

ĐẠI HỌC ĐÀ NẴNG  
TRƯỜNG ĐẠI HỌC BÁCH KHOA  
KHOA CÔNG NGHỆ THÔNG TIN

**ĐỒ ÁN TỐT NGHIỆP**  
NGÀNH: CÔNG NGHỆ THÔNG TIN  
CHUYÊN NGÀNH: Công nghệ thông tin

ĐỀ TÀI:

**XÂY DỰNG ỨNG DỤNG HỌC TIẾNG ANH  
THÔNG MINH VỚI CÔNG NGHỆ THỰC TẾ  
TĂNG CƯỜNG VÀ TRÍ TUỆ NHÂN TẠO**

Người hướng dẫn: ThS. NGUYỄN VĂN NGUYỄN  
Sinh viên thực hiện: NGUYỄN HỒNG TRUNG  
Mã số sinh viên: 102210283  
Lớp: 21TCLC\_DT4

Đà Nẵng, 06/2025

# TÓM TẮT

Tên đề tài: Xây dựng ứng dụng học tiếng Anh thông minh với công nghệ thực tế tăng cường và trí tuệ nhân tạo

Sinh viên thực hiện: Nguyễn Hồng Trung

Số thẻ SV: 102210283

Lớp: 21TCLC\_DT4

Trong bối cảnh hội nhập quốc tế, việc trang bị kỹ năng tiếng Anh cho trẻ em ngay từ nhỏ đang trở thành ưu tiên hàng đầu của nhiều bậc phụ huynh. Tuy nhiên, phần lớn các ứng dụng học tiếng Anh hiện nay chỉ dừng lại ở những bài học lặp đi lặp lại, thiếu yếu tố tương tác và khó tạo được hứng thú học tập lâu dài cho trẻ. Điều này dẫn đến việc trẻ dễ cảm thấy chán nản, thiếu động lực và không phát huy được tiềm năng ngôn ngữ của mình.

Để giải quyết vấn đề này, ứng dụng học tiếng Anh thông minh dành cho trẻ em, tích hợp công nghệ thực tế tăng cường (AR) và trí tuệ nhân tạo (AI) đã được phát triển. Ứng dụng này mang đến môi trường học tập sinh động kết hợp tương tác và các mô hình 3D, giúp trẻ không chỉ học từ vựng và phát âm mà còn rèn luyện kỹ năng giao tiếp. Công nghệ AR cho phép trẻ quan sát và tương tác với mô hình 3D ngay trong không gian thực tế, từ đó kích thích trí tò mò và tính khám phá, biến việc học thành một trải nghiệm thú vị.

Đồng thời, công nghệ AI sẽ đóng vai trò như một người bạn đồng hành thông minh, luôn sẵn sàng hướng dẫn, cung cấp lời khuyên và đề xuất bài học phù hợp với từng trẻ. Đảm bảo trẻ sẽ phát triển toàn diện cả về ngữ âm, từ vựng lẫn kỹ năng giao tiếp.

Báo cáo này được cấu trúc thành 5 phần chính:

- Mở đầu: Trình bày nhu cầu học tiếng Anh của trẻ em hiện nay và những khó khăn thường gặp trong phương pháp học tập truyền thống. Từ đó xác định mục tiêu và các tính năng chính của ứng dụng. Phân tích nhóm người dùng chính và phạm vi sử dụng, đồng thời giới thiệu tổng quan về công nghệ AR, AI và ngôn ngữ lập trình sử dụng.
- Chương 1: Cơ sở lý thuyết, trình bày lý thuyết các công nghệ sử dụng trong hệ thống.
- Chương 2: Phân tích nghiệp vụ và thiết kế hệ thống, cơ sở dữ liệu cũng như hệ thống tương tác AR/AI.
- Chương 3: Trình bày chi tiết quá trình triển khai hệ thống, mô tả chức năng nổi bật như tương tác với từ vựng, luyện nói, lưu giữ và xem tiến trình học tập.
- Kết luận: Trình bày kết quả đạt được, đánh giá kết quả, hạn chế và hướng phát triển.

## NHIỆM VỤ ĐỒ ÁN TỐT NGHIỆP

Họ tên sinh viên: Nguyễn Hồng Trung

Số thẻ sinh viên: 102210283

Lớp: 21TCLC\_DT4

Khoa: Công nghệ thông tin

Ngành: Kỹ thuật máy tính

1. Tên đề tài đồ án:

Xây dựng ứng dụng học tiếng Anh thông minh với công nghệ thực tế tăng cường và trí tuệ nhân tạo

2. Đề tài thuộc diện:  Có ký kết thỏa thuận sở hữu trí tuệ đối với kết quả thực hiện

3. Các số liệu và dữ liệu ban đầu:

Không có

4. Nội dung các phần thuyết minh và tính toán:

Nội dung của phần thuyết minh gồm:

**Mở đầu:** Nêu ra nhu cầu thực tế, đối tượng người dùng, giới thiệu đề tài, mục tiêu cần đạt được, các chức năng chính cần có trong hệ thống.

**Chương 1.** Cơ sở lý thuyết: Trình bày lý thuyết các công nghệ sử dụng và cách áp dụng vào hệ thống ứng dụng.

**Chương 2.** Phân tích và thiết kế hệ thống: Trình bày các sơ đồ, hướng giải quyết vấn đề

**Chương 3.** Triển khai và đánh giá kết quả: Trình bày cách thức thực hiện, vận hành, kiểm thử hệ thống và kết quả đạt được.

**Kết luận:** Đánh giá kết quả đạt được, hạn chế và hướng phát triển

5. Các bản vẽ, đồ thị (ghi rõ các loại và kích thước bản vẽ):

Không có.

6. Họ tên người hướng dẫn: ThS. Nguyễn Văn Nguyễn

7. Ngày giao nhiệm vụ đồ án: 9/4/2025

8. Ngày hoàn thành đồ án: ...../...../202.....

Đà Nẵng, ngày tháng 06 năm 2025

Trưởng Bộ môn .....

Người hướng dẫn

ThS. Nguyễn Văn Nguyễn

## LỜI NÓI ĐẦU

Xuyên suốt quá trình và thực hiện đề án tốt nghiệp tại trường Đại học Bách Khoa–Đại học Đà Nẵng, cũng như khoảng thời gian thực hiện và hoàn thành đề án này. Em đã nhận được sự hướng dẫn, giảng dạy nhiệt tình của quý Thầy, Cô trong khoa Công nghệ Thông Tin nói riêng và các khoa khác nói chung. Em xin gửi lời cảm ơn chân thành đến các thầy cô đã giảng dạy, hướng dẫn giúp em phát triển cả về kỹ năng lẫn kiến thức chuyên môn để em có được những kiến thức cần thiết để chuẩn bị cho tương lai của mình.

Đặc biệt, em xin gửi lời cảm ơn đến thầy Nguyễn Văn Nguyên đã hướng dẫn em trong suốt quá trình thực hiện đề án. Thầy là người đã dìu dắt, giúp em hiểu rõ mục tiêu đề án và đưa ra những lời khuyên về hướng phát triển cho đề tài. Thầy luôn dành thời gian gặp gỡ, thảo luận và giải đáp mọi thắc mắc của em.

Để có thể phát triển được đến ngày hôm nay, em xin bày tỏ lòng biết ơn đến gia đình, người thân, bạn bè đã luôn khích lệ, tạo cơ hội cho em tiếp tục học tập. Cảm ơn những người anh, chị đã luôn đồng hành và giúp em để em có thể hoàn thành tốt nhất quá trình học tập và thực hiện đề án của mình.

Tuy nhiên, trong học tập cũng như trong quá trình thực hiện vẫn sẽ không thể tránh khỏi những thiếu sót và hạn chế trong hệ thống. Em mong nhận được sự góp ý quý báu từ quý Thầy, Cô để em có thể hoàn thiện và đạt được thêm kinh nghiệm cho con đường phát triển sau này.

Em xin chân thành cảm ơn!

## CAM ĐOAN

Tôi xin cam đoan:

1. Báo cáo đồ án tốt nghiệp với tên đề tài “Xây dựng ứng dụng học tiếng Anh thông minh với công nghệ thực tế tăng cường và trí tuệ nhân tạo” là công trình nghiên cứu của chính cá nhân tôi dưới sự hướng dẫn trực tiếp của giảng viên ThS. Nguyễn Văn Nguyên.
2. Tôi đã tự đọc nghiên cứu, dịch tài liệu và tổng hợp các kiến thức đã làm nên báo cáo này và đảm bảo không sao chép ở bất cứ đâu.
3. Những lý thuyết trong luận văn đều được sử dụng tài liệu như tôi đã tham khảo ở phần tài liệu tham khảo đã có trong báo cáo.

Nếu có vi phạm, tôi xin chịu hoàn toàn trách nhiệm.

Sinh viên thực hiện

Nguyễn Hồng Trung

# MỤC LỤC

TÓM TẮT .....	3
NHIỆM VỤ ĐỒ ÁN TỐT NGHIỆP .....	4
LỜI NÓI ĐẦU.....	i
CAM ĐOAN .....	ii
MỤC LỤC .....	iii
DANH SÁCH CÁC BẢNG .....	v
DANH SÁCH CÁC HÌNH VẼ .....	vi
DANH SÁCH CÁC KÝ HIỆU, CHỮ VIẾT TẮT .....	viii
MỞ ĐẦU .....	1
CHƯƠNG 1: CƠ SỞ LÝ THUYẾT.....	5
1.1 Tổng quan về Unity và AR Foundation .....	5
1.1.1. Unity .....	5
1.1.2. AR Foundation.....	6
1.2. Thư viện ARCore – Android .....	7
1.3. AssetBundle trong Unity .....	8
1.4. Mô hình kiến trúc Client-Server .....	8
1.5. Tổng quan về Spring Boot và RESTful API .....	9
1.5.1. Spring Boot.....	9
1.5.2. RESTful API.....	9
1.6. Mô hình MVC .....	9
1.7. Hệ quản trị cơ sở dữ liệu MySQL .....	10
1.8. Tổng quan về FasterWhisper và thuật toán Levenshtein.....	10
1.8.1. FasterWhisper .....	10
1.8.2. Thuật toán Levenshtein.....	10
1.9. Ollama .....	10
1.10. LLaMA 3.2 .....	11
1.11. Công cụ hỗ trợ .....	11
1.12. Kết chương .....	11
CHƯƠNG 2: PHÂN TÍCH VÀ THIẾT KẾ HỆ THỐNG.....	12
2.1. Phân tích yêu cầu.....	12
2.1.1. Tổng quan.....	12
2.1.2. Chức năng của RESTful API Server.....	12
2.1.3. Chức năng của ứng dụng di động tích hợp AR .....	12
2.1.4. Yêu cầu phi chức năng.....	13
2.1.5. Yêu cầu đối với phần cứng .....	13

2.2. Phân tích và thiết kế hệ thống.....	14
2.2.1. Sơ đồ ca sử dụng.....	14
2.2.2. Sơ đồ tuần tự.....	20
2.2.3. Sơ đồ nguyên lý hoạt động .....	25
2.3. Thiết kế cơ sở dữ liệu .....	27
2.3.1. Cơ sở dữ liệu .....	27
2.3.2. Chi tiết các bảng trong cơ sở dữ liệu .....	28
2.4. Kết chương .....	30
CHƯƠNG 3: TRIỂN KHAI VÀ KẾT QUẢ ĐẠT ĐƯỢC.....	31
3.1. Server.....	31
3.1.1. Cài đặt các lớp truy xuất dữ liệu.....	31
3.1.2. Cài đặt các lớp xử lý logic.....	32
3.1.3. Cài đặt các lớp routing.....	32
3.1.4. Xác thực người dùng dùng sử dụng JWT .....	34
3.1.5. Sơ đồ xử lý các request.....	34
3.2. Ứng dụng di động.....	35
3.2.1. Lưu trữ mô hình và thông tin mô hình.....	36
3.2.2. Đăng nhập .....	37
3.2.3. Đăng ký.....	39
3.2.4. Màn hình chính.....	40
3.2.5. Chức năng “Xem thông tin” .....	41
3.2.6. Chức năng “Luyện phát âm” .....	45
3.2.7. Chức năng tương tác mô hình .....	49
3.2.8. Chức năng “Tiến độ học tập” .....	51
3.2.9. Chức năng “Trợ lý ảo” .....	53
3.3. Mô hình 3D.....	55
3.4. Phát âm từ vựng.....	55
3.5. Đánh giá kết quả .....	55
KẾT LUẬN VÀ HƯỚNG PHÁT TRIỂN.....	56
TÀI LIỆU THAM KHẢO .....	57

## **DANH SÁCH CÁC BẢNG**

Bảng 1.1: Các đối tượng chính trong XR Environment .....	6
Bảng 2.1: Đặc tả sơ đồ ca chức năng đăng nhập .....	15
Bảng 2.2: Đặc tả sơ đồ ca chức năng đăng ký .....	15
Bảng 2.3: Đặc tả sơ đồ ca chức năng đăng xuất .....	15
Bảng 2.4: Đặc tả sơ đồ ca chức năng quét hình ảnh hiển thị thông tin mô hình .....	16
Bảng 2.5: Đặc tả sơ đồ ca chức năng luyện phát âm kết hợp tương tác mô hình.....	17
Bảng 2.6: Đặc tả sơ đồ ca lưu trữ và xem tiến độ học tập.....	18
Bảng 2.7: Đặc tả sơ đồ ca chức năng quét trò chuyện với trợ lý ảo .....	19
Bảng 2.8: Bảng users .....	28
Bảng 2.9: Bảng models.....	28
Bảng 2.10: Bảng slides .....	28
Bảng 2.11: Bảng user_progress .....	29
Bảng 2.12: Bảng learning_progress .....	30
Bảng 3.1: Các lớp ánh xạ được sử dụng để truy xuất thông tin từ cơ sở dữ liệu .....	32
Bảng 3.2: Các lớp xử lý logic của hệ thống .....	32
Bảng 3.3: Danh sách các URL được lớp Controller quản lý .....	33
Bảng 3.4: Bảng danh sách các thao tác để tương tác với mô hình 3D .....	51

## DANH SÁCH CÁC HÌNH VẼ

Hình 1.1: Giao diện Unity Editor .....	5
Hình 1.2: Tính năng Image Tracking trong AR Foundation .....	7
Hình 1.3: Logo thư viện ARCore .....	7
Hình 1.4: Mô tả Asset Bundle .....	8
Hình 2.1: Sơ đồ ca sử dụng các chức năng tổng quát của người dùng.....	14
Hình 2.2: Sơ đồ ca sử dụng chức năng đăng nhập/ đăng ký/ đăng xuất.....	14
Hình 2.3: Sơ đồ ca sử dụng chức năng quét hình ảnh mô hình.....	16
Hình 2.4: Sơ đồ ca sử dụng chức năng luyện phát âm .....	17
Hình 2.5: Sơ đồ ca sử dụng chức năng lưu trữ và theo dõi tiến độ học tập.....	18
Hình 2.6: Sơ đồ ca sử dụng chức năng quản lý tài khoản .....	19
Hình 2.7: Sơ đồ tuần tự chức năng đăng nhập .....	20
Hình 2.8: Sơ đồ tuần tự chức năng đăng ký .....	20
Hình 2.9: Sơ đồ tuần tự chức năng “Thông tin mô hình”.....	21
Hình 2.10: Sơ đồ tuần tự chức năng “Luyện phát âm” .....	22
Hình 2.11: Sơ đồ tuần tự chức năng hiển thị bản trình chiếu thông tin mô hình .....	23
Hình 2.12: Sơ đồ tuần tự chức năng chuyển ngôn ngữ .....	23
Hình 2.13: Sơ đồ tuần tự chức năng “Tiến độ học tập” .....	24
Hình 2.14: Sơ đồ tuần tự chức năng “Trợ lý ảo” .....	24
Hình 2.15: Sơ đồ hoạt động chức năng đăng nhập/ đăng ký.....	25
Hình 2.16: Sơ đồ hoạt động chức năng “Thông tin mô hình” .....	25
Hình 2.17: Sơ đồ hoạt động chức năng “Luyện phát âm” .....	26
Hình 2.18: Sơ đồ hoạt động chức năng “Tiến độ học tập” và “Trợ lý ảo”.....	26
Hình 2.19: Sơ đồ thiết kế cơ sở dữ liệu .....	27
Hình 2.20: Mô tả các dữ liệu có trong các tệp Asset Bundle .....	27
Hình 3.1: Thiết kế kiến trúc server Spring Boot .....	31
Hình 3.2: Minh họa bảng models trong cơ sở dữ liệu và lớp Model ánh xạ .....	31
Hình 3.3: Sơ đồ xử lý request của server.....	35
Hình 3.4: Mô tả kiến trúc ứng dụng di động .....	36
Hình 3.5: Cửa sổ tạo AssetBundle với mô hình và thông tin mô hình.....	36
Hình 3.6: Cửa sổ để điều chỉnh mô hình gốc thành prefab cho AssetBundle.....	37
Hình 3.7: Biểu mẫu đăng nhập của ứng dụng .....	38
Hình 3.8: Sơ đồ xử lý đăng nhập.....	39
Hình 3.9: Biểu mẫu đăng ký của ứng dụng .....	39
Hình 3.10: Sơ đồ xử lý đăng ký.....	40
Hình 3.11: Giao diện màn hình trang chủ .....	40
Hình 3.12: Scene gốc chức năng “Xem thông tin”.....	41

Hình 3.13: Màn hình chờ tải mô hình sau khi chọn danh mục “animals” .....	41
Hình 3.14: Minh họa hiển thị mô hình trên hình ảnh được quét sau khi tải mô hình và quét hình ảnh thành công.....	42
Hình 3.15: Hiển thị thông tin mô hình khi nhấn vào mô hình.....	42
Hình 3.16: Sơ đồ xử lý tải mô hình và hiển thị mô hình .....	43
Hình 3.17: Sơ đồ xử lý hiển thị thông tin và phát âm thanh của mô hình.....	44
Hình 3.18: Sơ đồ xử lý hiển thị thông tin và phát âm thanh của mô hình.....	44
Hình 3.19: Giao diện hướng dẫn sử dụng chức năng luyện phát âm .....	45
Hình 3.20: Giao diện chức năng “Luyện phát âm” sau khi tải thành công mô hình .....	46
Hình 3.21: Hiển thị mô hình trên mặt phẳng đã được quét trên chức năng “Luyện phát âm” .	47
Hình 3.22: Tắt chức năng quét bề mặt phẳng.....	47
Hình 3.23: Minh họa chức năng đánh giá phát âm.....	48
Hình 3.24: Sơ đồ chức năng đánh giá phát âm.....	49
Hình 3.25: Thao tác xoay mô hình trái/phải, lên/xuống và thao tác phóng to/thu nhỏ mô hình .....	50
Hình 3.26: Sơ đồ hoạt động chức năng xoay phóng mô hình .....	50
Hình 3.27: Chức năng xem tiến độ học tập .....	52
Hình 3.28:Sơ đồ hoạt động chức năng xem tiến độ học tập.....	53
Hình 3.29: Chức năng “Trợ lý ảo” .....	54
Hình 3.30: Sơ đồ hoạt động chức năng “Trợ lý ảo” .....	54
Hình 3.31: Trang web “Pizza poly”.....	55
Hình 3.32: Trang web “TTS Maker”.....	55

## DANH SÁCH CÁC KÝ HIỆU, CHỮ VIẾT TẮT

Từ viết tắt	Diễn giải
CSDL	Cơ sở dữ liệu
CNTT	Công nghệ thông tin
AR	Augmented Reality
AI	Artificial Intelligence
API	Application Programing Interface
REST	Representational State Transfer
JSON	JavaScript Object Notation
2D/ 3D	2 dimensional/ 3 dimensional

## MỞ ĐẦU

### 1. Phân tích nhu cầu

Hiện nay, với xu hướng hội nhập và toàn cầu hóa đang khiến nhu cầu học ngoại ngữ tăng lên nhanh chóng. Cùng với đó là sự phát triển của các công nghệ mới đã tạo ra nhiều cơ hội, đặc biệt là trong lĩnh vực giáo dục. Việc ứng dụng công nghệ thực tế tăng cường (AR) và trí tuệ nhân tạo (AI) vào dạy tiếng Anh không chỉ mang lại trải nghiệm mới mẻ mà còn giúp nâng cao hiệu quả học tập cho trẻ em. Với sự phát triển của các nền tảng học tập tương tác thực tế, trẻ có thể dễ dàng tiếp cận bài học sinh động, trực quan và đầy thú vị. Đồng thời còn tạo động lực, kích thích sự ham học và khả năng sáng tạo của trẻ.

Dưới sự hỗ trợ của Trường Đại học Bách Khoa – Đại học Đà Nẵng và sự hướng dẫn tận tình của thầy Nguyễn Văn Nguyên, em đã phát triển đề tài “Ứng dụng học tiếng Anh thông minh với công nghệ thực tế tăng cường và trí tuệ nhân tạo”. Thông qua ứng dụng này, trẻ em có thể học từ vựng và luyện phát âm. Cùng với đó là khả năng tương tác với mô hình 3D thông qua công nghệ thực tế tăng cường sẽ tăng hứng thú và khả năng ghi nhớ từ vựng cho trẻ.

Với sự kết hợp giữa công nghệ hiện đại và nội dung học tập phù hợp, ứng dụng này sẽ mang lại trải nghiệm học tiếng Anh đặc biệt và đáng nhớ. Nó không chỉ giúp các em nhỏ phát triển vốn từ và kỹ năng phát âm, mà còn nâng cao sự sáng tạo và tính học hỏi – những yếu tố nền tảng để phát triển trong tương lai. Đặc biệt trong bối cảnh học trực tuyến và tự học tại nhà đang ngày càng phổ biến, một nền tảng học tiếng Anh sinh động, trực quan và mang tính cá nhân hóa cao sẽ là giải pháp hữu ích để nâng cao chất lượng giáo dục ngoại ngữ cho trẻ em Việt Nam.

Ứng dụng không chỉ đưa ra những bài học truyền thống mà còn hướng đến việc phát triển các bài học thành hệ sinh thái học tập toàn diện, nơi người học có thể trải nghiệm nhiều hình thức học khác nhau như các bản trình chiếu minh họa trực quan, mô hình tương tác thực tế, phản hồi giọng nói thời gian thực.

### 2. Mục đích

Mục đích của đề tài nhằm đến việc xây dựng một ứng dụng học tiếng Anh hiện đại, thân thiện, tích hợp công nghệ thực tế tăng cường (AR), nhằm hỗ trợ người học nâng cao kỹ năng ngôn ngữ một cách thú vị và hiệu quả. Ứng dụng không chỉ là một công cụ học tập truyền thống mà còn giúp người học tương tác với tài liệu học tập một cách linh hoạt, trao đổi kiến thức và thực hành ngôn ngữ một cách nhanh chóng, trực quan và hiệu quả.

Cụ thể, đề tài đặt ra các mục tiêu sau:

- Cung cấp các bài học theo chủ đề với nội dung đa dạng bao gồm từ vựng, phát âm.
- Cung cấp khả năng tương tác với mô hình: dịch chuyên, xoay, phóng to/ thu nhỏ
- Mô hình chứa các hoạt ảnh kích thích thị giác.
- Tăng cường khả năng ghi nhớ của người học thông qua các chức năng như mô hình 3D phát âm, slide trình chiếu minh họa, hiệu ứng hoạt họa đi kèm âm thanh, nhận diện hình ảnh thực tế để kích hoạt bài học tương ứng
- Tạo nền tảng cho việc phát triển các chức năng trong tương lai, như đồng bộ với sách giáo khoa hay sách thực tế tăng cường, tích hợp chatbot hỗ trợ học tập và phân tích dữ liệu học tập để nâng cao hiệu quả giảng dạy.

Ứng dụng hoạt động theo mô hình client-server, dữ liệu được lưu trữ trên mô hình, khi đó ứng dụng AR sẽ truy cập và tải xuống dữ liệu mới nhất về để sử dụng về sau.

Hệ thống sẽ gồm một RESTful API server đóng vai trò làm trung tâm lưu trữ dữ liệu. Nhận các request truy xuất, tạo, cập nhật, xóa dữ liệu và tiến hành xử lý, sau đó trả về response hoặc thông báo lỗi tương ứng.

AR mobile app là ứng dụng người dùng đầu cuối, truy xuất thông tin từ server thông qua API, xử lý hiển thị, thao tác với người dùng. Ứng dụng có thể sử dụng dữ liệu đã tải về trước đó (offline). Ứng dụng này sẽ có 4 chức năng chính là tương tác với mô hình, hiển thị bài học tương ứng với mô hình, luyện phát âm với mô hình, lưu trữ - xem tiến độ học tập và trò chuyện với trợ lý ảo

### 3. Tính năng

- Tương tác với mô hình 3D thông qua công nghệ AR: Khi quét hình ảnh hoặc chọn chủ đề, mô hình sẽ hiển thị trực tiếp trên không gian thật (camera), giúp học viên quan sát vật thể sinh động ở nhiều góc độ.
- Nghe phát âm mẫu từ mô hình: Khi nhấn vào mô hình 3D, hệ thống sẽ phát âm tên từ vựng tương ứng bằng tiếng Anh với giọng chuẩn (Anh/Mỹ) hoặc tiếng Việt.
- Hiển thị slide minh họa thông tin: Sau khi tương tác với mô hình, hệ thống sẽ hiển thị thông tin bổ sung như hình ảnh, thông tin về mô hình, nghĩa tiếng Việt, câu chuyện liên quan đến từ vựng.
- Luyện phát âm và chấm điểm: Người dùng có thể thu âm phát âm của mình, hệ thống sẽ phân tích, chấm điểm và phản hồi giúp cải thiện kỹ năng nói
- Xem tiến trình học tập cá nhân: Học viên có thể theo dõi các chủ đề đã học, số điểm luyện phát âm, số từ đã hoàn thành và tiến độ tổng thể.

### 4. Những nghiên cứu liên quan

#### 4.1. Ứng dụng *Delightex*

Ứng dụng **Delightex** hỗ trợ nhiều nền tảng như Android, iOS, và cả trình duyệt web, giúp người dùng tạo nội dung từ mọi nơi. Trong môi trường AR, người học có thể đặt và tương tác với các đối tượng 3D ngay trong không gian thực, giúp tăng tính nhập vai và trực quan trong quá trình học tập. Delightex tập trung vào việc tích hợp lập trình (coding), tư duy thiết kế, và khả năng sáng tạo vào môi trường học tập tương tác.

CoSpaces Edu được thiết kế cho giáo dục, cho phép người dùng không chỉ đặt từng đối tượng đơn lẻ mà còn có thể lập trình chuyển động, tương tác và xây dựng cả một kịch bản học tập đa phương tiện trong không gian ảo. Tuy nhiên ứng dụng không thực sự hỗ trợ học tiếng Anh.



Hình 1. Ứng dụng Delightex trên CH Play

#### 4. Đối tượng và phạm vi nghiên cứu

##### 4.1 Đối tượng nghiên cứu

Đối tượng nghiên cứu chính của đề tài là sử dụng thư viện lập trình “ARCore SDK for Unity” và “AR Foundation” – gói hỗ trợ các giao diện lập trình các chức năng AR để phát triển một ứng dụng hỗ trợ học tiếng Anh trên hệ điều hành Android. Cùng với đó là nghiên cứu về việc xây dựng một kiến trúc hệ thống sử dụng máy chủ để lưu trữ - xử lý dữ liệu tương tác với ứng dụng di động. Cuối cùng

là nghiên cứu cách tích hợp hệ thống nhận diện giọng nói FasterWhisper và mô hình ngôn ngữ lớn LLaMA3.2 hỗ trợ hỏi – đáp để tăng tính tương tác cho ứng dụng. Đối tượng nghiên cứu đầy đủ sẽ là một hệ thống gồm 2 thành phần:

- RESTful API Server: Xây dựng một Server cung cấp các thao tác truyền tải các tệp mô hình 3D, tiến độ học tập, thông tin mô hình, xử lý các API nhận diện giọng nói, chấm điểm phát âm, phản hồi kết quả từ mô hình ngôn ngữ lớn cho các yêu cầu của ứng dụng di động
- AR Mobile App: Phát triển ứng dụng tích hợp công nghệ AR trên hệ điều hành Android, cho phép truy xuất – giao tiếp với server, lưu trữ và sử dụng dữ liệu. Ứng dụng này là ứng dụng cốt lõi của hệ thống, giúp người dùng mang trải nghiệm AR vào chương trình học tập

#### 4.2 Phạm vi nghiên cứu

Phạm vi nghiên cứu của đề tài xoay quanh việc xử lý các thành phần, mã nguồn, mô hình trên phần mềm game engine Unity 3D kết hợp với thư gói hỗ trợ AR Foundation – thư viện AR Core để tạo ra trải nghiệm AR. Ngoài ra còn nghiên cứu về một vài thư viện như JWT API, Spring Boot framework, Spring Security và các thư viện hỗ trợ để xây dựng một hệ thống RESTful API hỗ trợ quản lý thông tin. Cuối cùng là cách ứng dụng FasterWhisper và mô hình LLaMA 3.2 vào server.

### 5. Phương pháp nghiên cứu

Nội dung mà đề tài hướng đến chưa được phổ biến rộng rãi, vì vậy phương pháp nghiên cứu chủ yếu là tìm hiểu thông qua các tài liệu trên trang chủ của các thư viện như AR Core – Google, AR Foundation – Unity. Cùng với đó là những video trên Youtube. Từ đó vận dụng, kết hợp những kiến thức đã đạt được vào việc phát triển đề tài. Đối với mỗi thành phần của hệ thống sẽ có những phương pháp nghiên cứu khác nhau.

Spring Boot Server: Nghiên cứu về các thư viện hỗ trợ, cách xây dựng một hệ thống RESTful Framework, Spring Security với JWT token – authentication, cách giao tiếp với các mô hình học máy và cách liên kết với hệ quản trị cơ sở dữ liệu MySQL.

AR Mobile App: Nghiên cứu về ứng dụng Unity 3D, gói hỗ trợ AR Foundation, cách truyền và xử lý dữ liệu với server, cách tương tác với mô hình 3D trong không gian thực.

### 6. Công nghệ phát triển

- Ngôn ngữ: Java, Python, C#
- Framework/Library: Spring Boot, Unity AR Foundation
- Database: MySQL
- Mã hóa & bảo mật: Spring Security, Json Web Token

## CHƯƠNG 1: CƠ SỞ LÝ THUYẾT

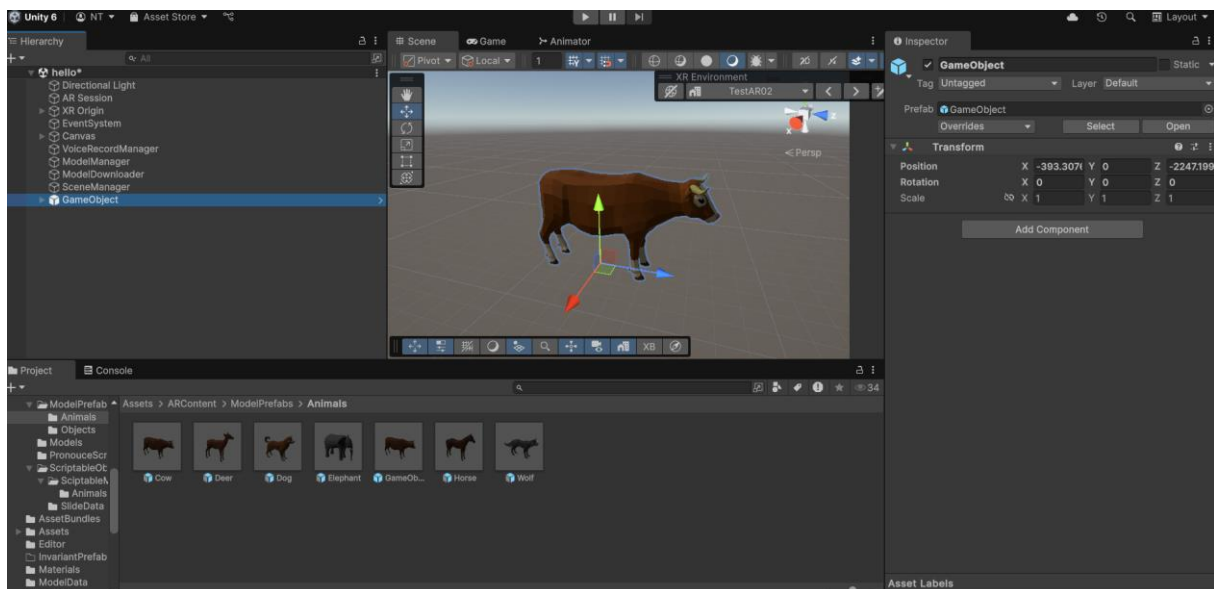
### 1.1 Tổng quan về Unity và AR Foundation

#### 1.1.1. Unity

Unity là một công cụ phát triển ứng dụng game và tương tác 3D phổ biến, hỗ trợ đa nền tảng như Windows, Android, iOS, WebGL. Unity sử dụng ngôn ngữ lập trình C#, nền tảng này cho phép xây dựng các ứng dụng có tính tương tác cao với giao diện trực quan, hỗ trợ hiển thị, cấu hình và điều khiển các đối tượng 3D linh hoạt. Unity ban đầu được thiết kế dành cho phát triển trò chơi, nhưng ngày nay đã được ứng dụng rộng rãi trong nhiều lĩnh vực khác nhau như giáo dục, kiến trúc và thực tế ảo/tăng cường. Trong dự án này, Unity được sử dụng để xây dựng giao diện người dùng trên thiết bị di động, hiển thị mô hình 3D minh họa từ vựng tiếng Anh và tạo môi trường học tập tương tác thực tế thông qua công nghệ thực tế tăng cường.

Các tính năng nổi bật của Unity bao gồm:

- Hỗ trợ hệ thống vật lý, ánh sáng, hiệu ứng hình ảnh và hoạt ảnh.
- Cộng đồng phát triển lớn và kho tài nguyên Asset Store phong phú.
- Tích hợp dễ dàng với các plugin, dịch vụ bên thứ ba.
- Khả năng dựng và tương tác với các đối tượng 3D trong môi trường mô phỏng thực tế.



Hình 1.1: Giao diện Unity Editor

Trong Unity còn có những công cụ nhỏ khác giúp phục vụ cho các mục đích khác nhau trong quá trình xây dựng ứng dụng. Cụ thể dưới đây là một số đối tượng thường được sử dụng trong quá trình phát triển ứng dụng với Unity Engine:

- GameObject – Đối tượng cơ bản nhất
- Component – Thành phần mở rộng
- Prefab – Mẫu đối tượng tái sử dụng

- Scene – Môi trường làm việc
- Asset – Tài nguyên
- ScriptableObject – Dữ liệu cấu hình
- MonoBehaviour – Lớp lập trình hành vi
- Camera và Light
- XR Environment – Môi trường Thực tế mở rộng (AR/VR/MR)
  - XR (Extended Reality) bao gồm:
    - AR (Augmented Reality): Hiện thị đối tượng ảo trên thế giới thực.
    - VR (Virtual Reality): Tạo môi trường ảo hoàn toàn.
    - MR (Mixed Reality): Kết hợp giữa AR và VR, có tương tác vật lý giữa thế giới ảo và thực.
  - Các đối tượng chính trong XR:

*Bảng 1.1: Các đối tượng chính trong XR Environment*

Object	Mô tả
AR Session	Quản lý phiên AR (gồm tracking, ánh sáng...)
AR Session Origin	Chứa camera và các đối tượng để hiển thị AR
AR Camera / XR Camera	Camera dùng cho AR hoặc VR, có tracking chuyển động đầu
AR Plane Manager	Phát hiện mặt phẳng trong AR
AR Raycast Manager	Xử lý tương tác với bề mặt qua raycast
XR Interaction Toolkit	Bộ công cụ hỗ trợ tương tác tay, vật lý, UI cho XR

### 1.1.2. AR Foundation

AR là viết tắt của Augmented Reality – Thực tế tăng cường. Nó là sự tích hợp thông tin kỹ thuật số với môi trường thực tế của người dùng trong thời gian thực. Hay với cách nói khác, môi trường tồn tại xung quanh được che phủ bởi thông tin kỹ thuật số để tạo cho người dùng trải nghiệm như đang tương tác với những thông tin đó

AR Foundation là một gói hỗ trợ xây dựng ứng dụng thực tế tăng cường do Unity phát triển, cho phép xây dựng các ứng dụng AR có thể hoạt động trên nhiều nền tảng như ARKit (iOS) và ARCore (Android). Đây là lớp trung gian giúp các lập trình viên không cần phải làm việc trực tiếp với các SDK nền tảng gốc mà vẫn có thể truy cập và sử dụng các tính năng AR phổ biến.

Các chức năng chính của AR Foundation:

- Phát hiện mặt phẳng (plane detection)
- Raycasting
- Truy vết chuyển động (motion tracking)
- Hiện thị đối tượng ảo
- Hiệu ánh sáng thực (light estimation)

- Nhận diện hình ảnh, khuôn mặt

Cấu trúc chính của AR Foundation

- AR Session
- AR Session Origin
- AR Camera

Trong ứng dụng này, AR Foundation đóng vai trò quan trọng trong việc tích hợp và hiển thị mô hình 3D các đối tượng liên quan đến từ vựng như con vật, đồ vật, phương tiện trong môi trường thực tế của học viên.



*Hình 1.2: Tính năng Image Tracking trong AR Foundation*

## **1.2. Thư viện ARCore – Android**

Gói hỗ trợ phát triển ứng dụng AR Foundation chứa những giao diện lập trình cho các chức năng AR, nhưng nó không bao gồm bất kỳ chức năng nào cả. Để sử dụng AR Foundation trên một nền tảng cụ thể, thiết bị cần phải cài đặt một gói plug-in tương ứng cho thiết bị đó. Ở đây, plug-in cần thiết để thiết bị Android xây dựng và chạy ứng dụng là ARCore Plug-in.



*Hình 1.3: Logo thư viện ARCore*

ARCore là một bộ thư viện phát triển ứng dụng do Google phát triển. ARCore giúp hỗ trợ xây dựng những ứng dụng tích hợp thực tế tăng cường. Với những API của

ARCore, thiết bị di động sẽ có khả năng cảm nhận và hiểu được môi trường xung quanh, có khả năng tương tác với thông tin thu được

Công nghệ chính của ARCore để tích hợp thông tin kỹ thuật số vào thế giới thật:

- **Motion-tracking:** theo dõi, đánh dấu vị trí tương đối của thiết bị so với thế giới thật
- **Environment understanding:** cho phép điện thoại phát hiện vị trí của những mặt phẳng như mặt đất, mặt sàn, mặt bàn, tường.
- **Light estimation:** nhận diện thông tin về ánh sáng của môi trường xung quanh và cung cấp độ đậm nhạt, màu sắc tương ứng cho thông tin kỹ thuật số.
- **User interaction:** ARCore sử dụng hàm kiểm tra điểm nhấn. Điều này cho phép người dùng chọn – tương tác với vật thể ảo trong môi trường thực tế.

Các thiết bị hỗ trợ ARCore được liệt kê tại trang:

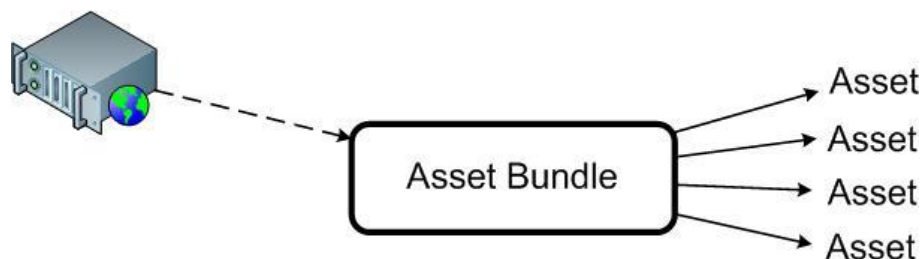
<https://developers.google.com/ar/devices>

### 1.3. AssetBundle trong Unity

Asset Bundle là một tập hợp dữ liệu tài nguyên (asset) được biên dịch và đóng gói bởi Unity. Các asset này có thể là:

- Mô hình 3D
- Âm thanh, nhạc nền
- Hình ảnh, texture
- Animation
- Prefab
- Scene

Asset Bundle được lưu dưới dạng file nhị phân (binary) và có thể được tải vào trong game tại thời điểm **runtime** (khi ứng dụng đang chạy), thay vì nhúng sẵn vào file cài đặt.



Hình 1.4: Mô tả Asset Bundle

### 1.4. Mô hình kiến trúc Client-Server

Ứng dụng được xây dựng theo mô hình Client-Server, gồm 2 thành phần chính:

- Client:
  - Là ứng dụng được phát triển bằng Unity và cài đặt trên thiết bị người dùng (Android)
  - Giao diện người dùng (UI) hiển thị các chức năng chính

- Unity sử dụng các HTTP request để giao tiếp với server thông qua các API RESTful
- Dữ liệu trao đổi gồm thông tin người dùng, danh sách mô hình, danh sách thông tin mô hình, tệp âm thanh, tiến trình học.
- Server:
  - Ứng dụng web backend được phát triển dựa trên framework Spring Boot
  - Chịu trách nhiệm xử lý các yêu cầu từ phía client Unity, tương tác với cơ sở dữ liệu và thực hiện logic nghiệp vụ
  - Cung cấp các REST API endpoint để Unity có thể đăng nhập, đăng ký người dùng, lấy danh sách mô hình 3D, nhận và lưu trữ file, trả về kết quả xử lý.
  - Các API sử dụng định dạng JSON để trao đổi dữ liệu, tệp đính kèm sử dụng multipart/form-data

## 1.5. Tổng quan về Spring Boot và RESTful API

### 1.5.1. Spring Boot

Spring Boot là một Framework back-end mạnh mẽ và linh hoạt, nó là một phần của hệ sinh thái Spring, được phát triển nhằm đơn giản hóa quá trình xây dựng hệ thống back-end với Java. Ưu điểm chính của Spring Boot là cấu hình tối giản, tích hợp sẵn các thành phần cần thiết giúp nhanh chóng phát triển ứng dụng RESTful

Trong dự án này, Spring Boot đảm nhiệm vai trò xây dựng hệ thống back-end, cụ thể:

- Xác thực và phân quyền truy cập
- Cung cấp API cho ứng dụng Unity lấy mô hình 3D, âm thanh, thông tin các bản trình chiếu của từ vựng
- Nhận và xử lý tệp âm thanh gửi từ client để chấm điểm phát âm
- Ghi nhận tiến trình học tập của người dùng

### 1.5.2. RESTful API

RESTful API là một chuẩn kiến trúc thiết kế các API giúp các ứng dụng client (Unity) và server (Spring Boot) trao đổi dữ liệu qua giao thức HTTP. Mỗi endpoint của REST API tương ứng với một tài nguyên cụ thể (từ vựng, mô hình 3D, tệp âm thanh)

Trong dự án, Unity gọi các REST API để tải nội dung học tập, gửi tệp ghi âm, nhận đánh giá, quản lý tài khoản học viên và đồng bộ dữ liệu học tập

## 1.6. Mô hình MVC

Mô hình MVC là một kiến trúc phổ biến được sử dụng trong các ứng dụng web, trong đó

- Model đại diện cho dữ liệu và logic nghiệp vụ (lớp Model, lớp User,...)
- View là giao diện người dùng (trong hệ thống này, View được thể hiện qua ứng dụng Unity)
- Controller là lớp trung gian điều phối yêu cầu, xử lý dữ liệu từ Model và trả về cho View

Trong Spring Boot, mô hình MVC được triển khai qua các lớp:

- @Controller hoặc @RestController để xử lý yêu cầu
- @Service để viết logic nghiệp vụ
- @Repository để giao tiếp với cơ sở dữ liệu

Mô hình này giúp hệ thống dễ bảo trì, mở rộng và tách biệt rõ ràng giữa các thành phần và tăng tính tái sử dụng.

## 1.7. Hệ quản trị cơ sở dữ liệu MySQL

MySQL là hệ quản trị cơ sở dữ liệu quan hệ mã nguồn mở được sử dụng rộng rãi trong phát triển ứng dụng web và doanh nghiệp. MySQL hoạt động theo mô hình client-server, hỗ trợ ngôn ngữ truy vấn SQL để tạo, truy xuất, cập nhật và quản lý dữ liệu.

Trong dự án này, MySQL được dùng để lưu trữ thông tin người học, tiến trình học tập, từ vựng, mô hình 3D và lịch sử phát âm, từ đó hỗ trợ việc truy xuất và phân tích dữ liệu học tập một cách hiệu quả.

## 1.8. Tổng quan về FasterWhisper và thuật toán Levenshtein

### 1.8.1. FasterWhisper

FasterWhisper là một phiên bản được tối ưu hóa về hiệu suất của mô hình Whisper do OpenAI phát triển, có khả năng nhận diện giọng nói chính xác với độ trễ thấp. Trong hệ thống học tiếng Anh, FasterWhisper đảm nhiệm vai trò

- Phân tích tệp ghi âm gửi từ ứng dụng Unity
- Chuyển âm thanh thành văn bản

FasterWhisper giúp việc đánh giá phát âm trở nên chính xác và khách quan hơn, tạo điều kiện để trẻ tự luyện phát âm hiệu quả

### 1.8.2. Thuật toán Levenshtein

Levenshtein là một thuật toán tính khoảng cách giữa hai chuỗi, dựa trên số bước chèn, xóa hoặc thay ký tự để biến chuỗi A thành chuỗi B. Trong hệ thống này, thuật toán dùng để so sánh kết quả giọng nói đã chuyển thành văn bản với từ vựng chuẩn

- Nếu hai chuỗi giống nhau: phát âm chính xác
- Khoảng cách càng nhỏ: phát âm càng gần đúng
- Dựa trên tỉ lệ sai lệch để phân loại và chấm điểm: Chính xác, Chưa chính xác

## 1.9. Ollama

Tạo ra nền tảng dễ sử dụng để chạy mô hình ngôn ngữ lớn (LLM) trên máy cá nhân, server cục bộ, hoặc cloud riêng

- Local-first AI: Thay vì gửi dữ liệu đến server đám mây (như OpenAI), Ollama nhắm đến việc chạy mô hình hoàn toàn tại máy local, giúp bảo mật và giảm độ trễ.
- Quantization: Áp dụng kỹ thuật giảm độ chính xác của trọng số (từ FP16 xuống INT4/INT8) để giảm kích thước mô hình, giúp nó chạy trên CPU/GPU yếu mà vẫn giữ hiệu suất tương đối.

- Memory-efficient inference: Sử dụng backend như llama.cpp (sử dụng thư viện GGML) giúp suy luận hiệu quả bằng CPU, rất phù hợp cho desktop/server không có GPU mạnh.
- Truy cập API chuẩn OpenAI: Ollama expose ra API theo chuẩn OpenAI Chat API (/v1/chat/completions) => dễ tích hợp.

### **1.10. LLaMA 3.2**

LLaMA 3.2 là một mô hình ngôn ngữ lớn (LLM) được cập nhật cải tiến từ mô hình LLaMA 3. Mô hình LLaMA dựa trên kiến trúc decoder-only Transformer, tương tự GPT. Mô hình phiên bản 3.2 có những áp dụng cải tiến:

- Rotary positional embedding (RoPE)
- Group Query Attention (GQA): Tối ưu attention với nhiều head.
- SwiGLU activation
- Normalization improvements (RMSNorm hoặc variant)

Tập huấn luyện:

- Huấn luyện trên 1+ triệu token (tương đương hàng trăm tỷ từ).
- Dữ liệu đa lĩnh vực: web, khoa học, văn học, hội thoại.
- Một số mô hình được Instruction-tuned (ví dụ: LLaMA-3-Instruct) giúp phản hồi theo yêu cầu tự nhiên.

### **1.11. Công cụ hỗ trợ**

Các công cụ bên ngoài được sử dụng để hỗ trợ xây dựng hệ thống bao gồm:

- Trello: Theo dõi tiến độ dự án.
- Postman: Kiểm thử và tài liệu hóa API
- Ngrok: Công cụ tạo đường hầm giúp chạy demo dự án.

### **1.12. Kết chương**

Trong chương này đã trình bày cơ sở lý thuyết của các công nghệ chính được sử dụng trong dự án bao gồm Unity, AR Foundation, Spring Boot, REST API, mô hình MVC, MySQL, cùng với các công cụ AI như FasterWhisper và thuật toán Levenshtein. Việc nắm rõ những công nghệ này giúp đảm bảo ứng dụng được xây dựng trên nền tảng vững chắc, dễ mở rộng và đáp ứng tốt các yêu cầu về tính năng, hiệu suất và trải nghiệm người dùng.

Chương tiếp theo sẽ trình bày kiến trúc hệ thống tổng thể, sơ đồ tương tác giữa client (Unity) và server (Spring Boot), cơ sở dữ liệu người dùng và thiết kế các luồng hoạt động trong hệ thống.

## CHƯƠNG 2: PHÂN TÍCH VÀ THIẾT KẾ HỆ THỐNG

### 2.1. Phân tích yêu cầu

#### 2.1.1. Tổng quan

Ứng dụng hỗ trợ học tiếng Anh tích hợp công nghệ thực tế tăng cường và trí tuệ nhân tạo được xây dựng nhằm nâng cao trải nghiệm học tập của người dùng, gia tăng sự hứng thú của trẻ em đối với ngôn ngữ. Mục tiêu của ứng dụng là cung cấp một môi trường học tập sinh động và trực quan. Để đáp ứng những yêu cầu đó, hệ thống phải đáp ứng những nhiều cơ bản sau:

- Hỗ trợ tương tác giữa người dùng và mô hình 3D
- Giao diện tối giản phù hợp, thân thiện với trẻ em.
- Đảm bảo tính bảo mật tài khoản

#### 2.1.2. Chức năng của RESTful API Server

Xác thực người dùng với JWT Token: tương ứng với mỗi request được gửi tới, hệ thống sẽ kiểm tra token được gửi kèm với header để xác thực người dùng.

- **Quản lý dữ liệu mô hình 3D:** lưu trữ mô hình trên hệ thống tập tin, thông tin mô hình trên cơ sở dữ liệu, nhận các request yêu cầu mô hình kèm thông tin và trả về kết quả.
- **Trung gian cho hệ thống nhận diện giọng nói, hệ thống chatbot, chấm điểm phát âm:** trung gian giữa người dùng và hệ thống nhận diện giọng nói FasterWhisper, cũng là trung gian giữa người dùng và hệ thống quản lý mô hình học máy Ollama với mô hình ngôn ngữ lớn LLaMA 3.2
- **Quản lý thông tin các bài học liên quan tới mô hình:** lưu trữ, xử lý thông tin các bài học, câu chuyện liên quan tới mô hình trên hệ thống tập tin, cơ sở dữ liệu.
- **Quản lý thông tin người dùng dữ liệu học tập:** lưu trữ, quản lý dữ liệu học tập của từng người dùng. Nhận các request yêu cầu trả về tiến độ học tập.

#### 2.1.3. Chức năng của ứng dụng di động tích hợp AR

- **Đăng ký tài khoản, đăng nhập hệ thống:** đăng nhập, đăng ký, lưu thông tin người dùng, JWT Token để thực hiện các request khác
- **Chọn chủ đề học theo nhu cầu cá nhân:** với mỗi chủ đề được chọn, gửi request tải các mô hình và thông tin liên quan đến mô hình đó thuộc chủ đề được chọn, lưu trữ vào bộ nhớ tạm để sử dụng sau này.
- **Hiển thị mô hình 3D sinh động và rõ ràng để minh họa từ vựng:** hiển thị trên bề mặt hình ảnh được quét, hoặc hiển thị trên bề mặt phẳng được phát hiện
- **Tương tác với mô hình được hiển thị:** cho phép người dùng tương tác, xoay phóng mô hình.
- **Quản lý hiển thị mặt phẳng được phát hiện:** cho phép người dùng tùy chọn ẩn hoặc hiện mặt phẳng đã được phát hiện bởi AR Session.

- **Hiển thị bản trình chiếu chứa thông tin liên quan tới mô hình 3D được chọn:** hiển thị thông tin, câu chuyện liên quan tới mô hình 3D theo dạng chữ viết và hình ảnh.
- **Phát âm thanh phát âm mẫu và ghi âm giọng nói của người đọc:** cho phép người dùng thực hiện luyện phát âm từ vựng và xem được điểm phát âm.
- **Đồng bộ tiến trình học lên máy chủ:** với mỗi lần luyện phát âm và xem các bản thuyết trình, tiến độ học tập sẽ được đồng bộ hóa lên máy chủ
- **Quản lý, bảo mật tài khoản học viên và thông tin cá nhân**

#### ***2.1.4. Yêu cầu phi chức năng***

Ngoài ra, hệ thống còn phải đáp ứng một số yêu cầu phi chức năng nhằm đảm bảo tính bảo mật và duy trì khả năng mở rộng trong quá trình phát triển về sau

- Tính bảo mật
  - Thông tin người dùng phải được lưu trữ an toàn
  - Sử dụng chuẩn HTTPS, đảm bảo dữ liệu không rò rỉ khi truyền tải.
  - Mật khẩu được mã hóa sử dụng thuật toán HS512
- Khả năng mở rộng
  - Hệ thống được thiết kế theo kiến trúc Client – Server, không đính kèm toàn bộ mô hình lên ứng dụng Unity, cho phép mở rộng từng phần, thêm chức năng, danh mục khi cần thiết
- Dễ sử dụng
  - Giao diện người dùng được thiết kế đơn giản, phù hợp với trẻ em
  - Các chức năng chính được bố trí rõ ràng, dễ tiếp cận

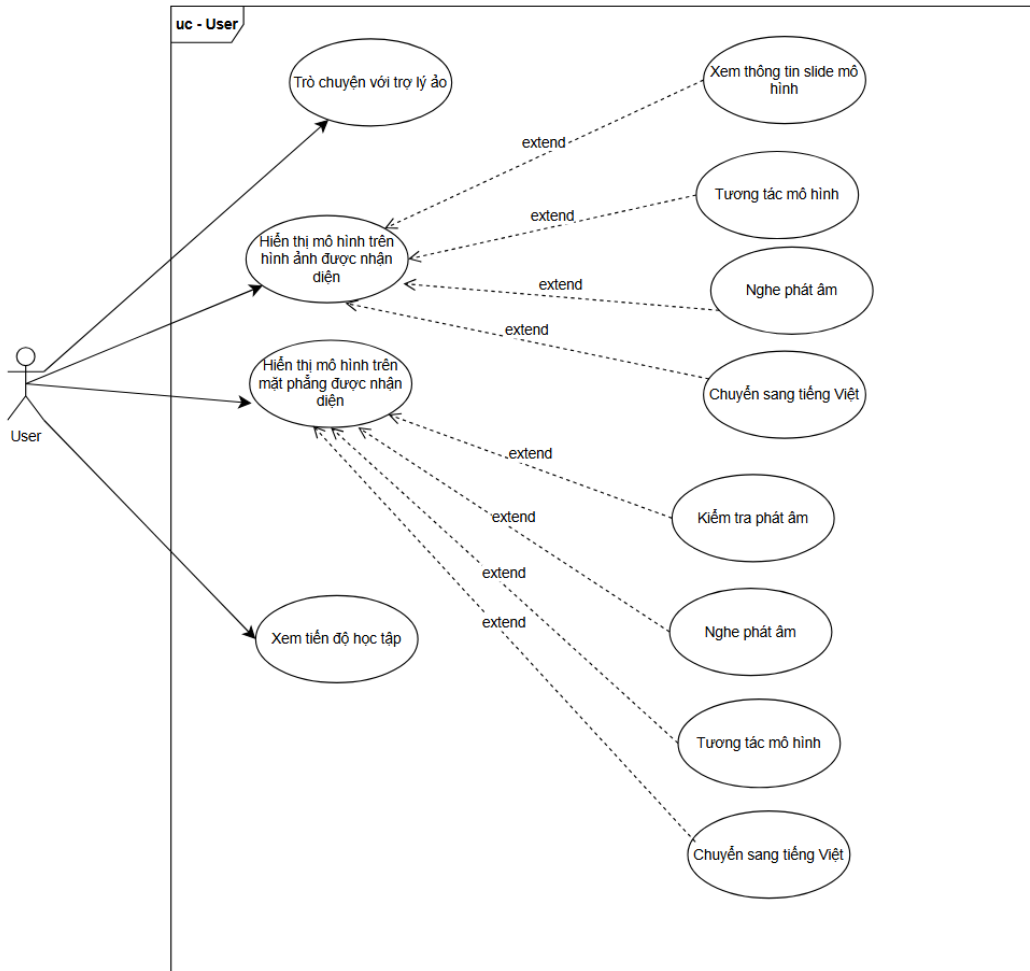
#### ***2.1.5. Yêu cầu đối với phần cứng***

Ứng dụng chạy trên các thiết bị di động với hệ điều hành Android 7.0'Nougat' trở lên và có hỗ trợ AR Core, camera, microphone.

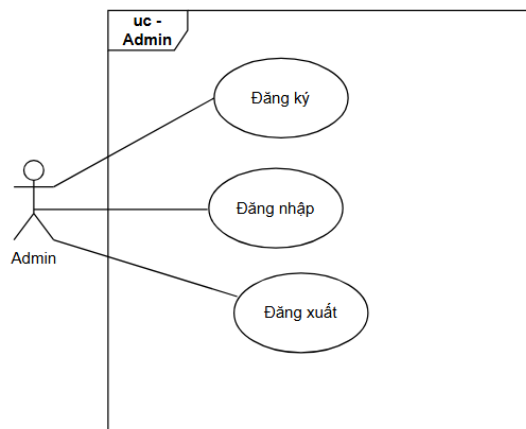
## 2.2. Phân tích và thiết kế hệ thống

### 2.2.1. Sơ đồ ca sử dụng

Sơ đồ ca sử dụng của người dùng: các chức năng người dùng có thể sử dụng trong hệ thống



Hình 2.1: Sơ đồ ca sử dụng các chức năng tổng quát của người dùng



Hình 2.2: Sơ đồ ca sử dụng chức năng đăng nhập/ đăng ký/ đăng xuất

*Bảng 2.1: Đặc tả sơ đồ ca chức năng đăng nhập*

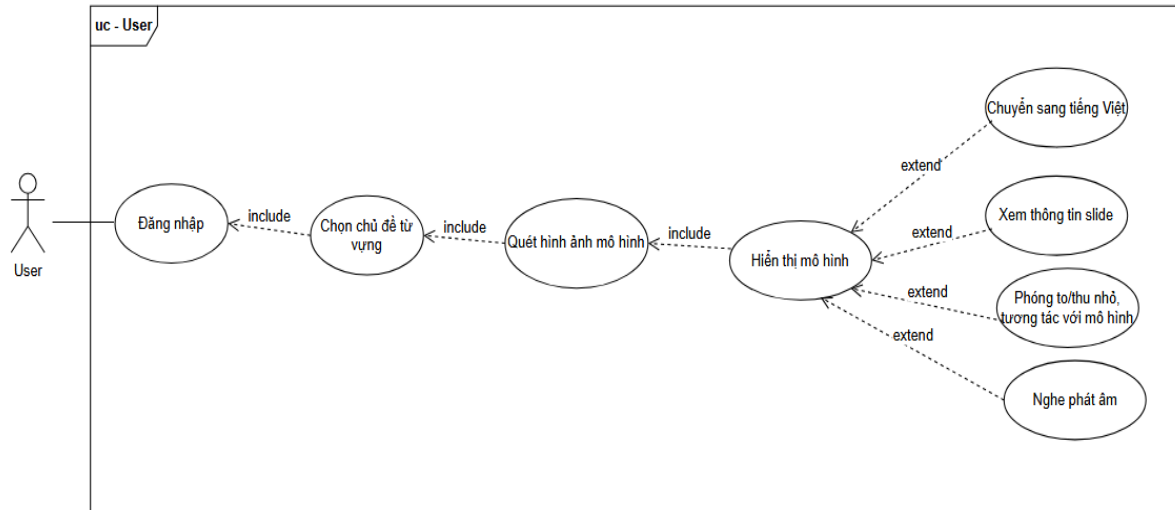
Tên Usecase	Người dùng đăng nhập
Tác nhân chính	User
Mô tả usecase	Người dùng đăng nhập bằng tên tài khoản, mật khẩu để truy cập vào hệ thống
Tiền điều kiện	Chưa đăng nhập. Đã đăng kí tài khoản.
Hậu điều kiện	Đăng nhập thành công, vào trang chủ.
Kích hoạt	Người dùng nhấn nút “Đăng nhập” trên màn hình khởi đầu.
<p>Chuỗi sự kiện</p> <ol style="list-style-type: none"> <li>1. Người dùng nhập tên tài khoản, mật khẩu và nhấn nút “Đăng nhập”</li> <li>2. Ứng dụng Unity kiểm tra thông tin nhập vào. Nếu hợp lệ sẽ gửi sang server, nếu server kiểm tra thông tin có trong hệ thống sẽ trả về token, Unity nhận token, lưu vào bộ nhớ cục bộ và chuyển hướng người dùng đến màn hình chính. Nếu thông tin đăng nhập không hợp lệ sẽ báo lỗi</li> </ol>	

*Bảng 2.2: Đặc tả sơ đồ ca chức năng đăng ký*

Tên Usecase	Người dùng đăng ký
Tác nhân chính	User
Mô tả usecase	Người dùng đăng ký tài khoản
Tiền điều kiện	Chưa có tài khoản
Hậu điều kiện	Đăng ký thành công, chuyển hướng về trang đăng nhập
Kích hoạt	Người dùng nhấn nút “Đăng ký” trên màn hình khởi đầu.
<p>Chuỗi sự kiện</p> <ol style="list-style-type: none"> <li>1. Người dùng nhập tên tài khoản, mật khẩu, email và nhấn nút “Đăng ký”</li> <li>2. Ứng dụng Unity kiểm tra thông tin nhập vào. Nếu hợp lệ sẽ gửi sang server, nếu server kiểm tra thông tin có trong hệ thống sẽ trả về token, Unity nhận token, lưu vào bộ nhớ cục bộ và chuyển hướng người dùng đến trang đăng nhập. Nếu thông tin đăng nhập không hợp lệ sẽ báo lỗi</li> </ol>	

*Bảng 2.3: Đặc tả sơ đồ ca chức năng đăng xuất*

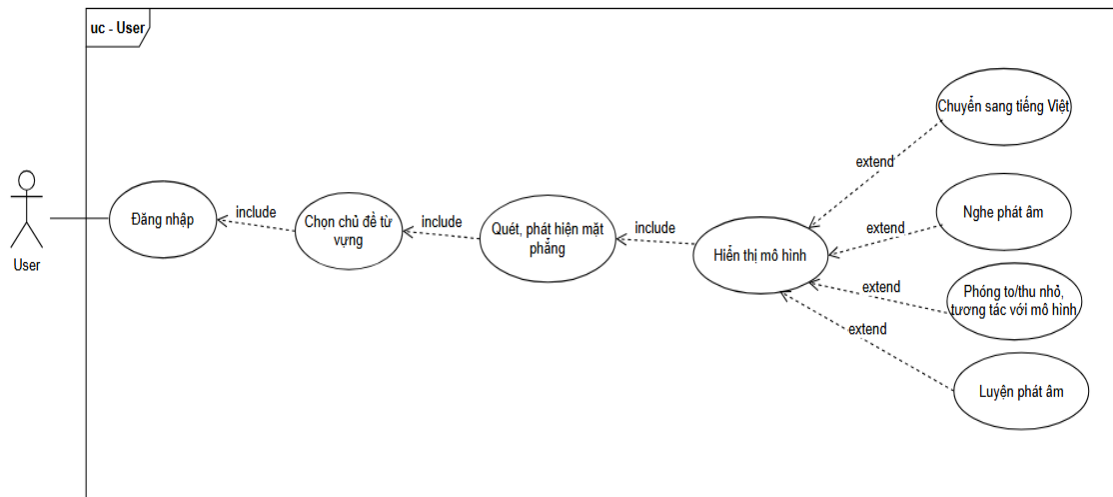
Tên Usecase	Người dùng đăng xuất
Tác nhân chính	User
Mô tả usecase	Người dùng đăng xuất tài khoản
Tiền điều kiện	Đã đăng nhập
Hậu điều kiện	Chuyển hướng về trang đăng nhập
Kích hoạt	Người dùng nhấn nút “Đăng xuất” trên màn hình chính.
<p>Chuỗi sự kiện</p> <ol style="list-style-type: none"> <li>1. Người dùng nhấn nút “Đăng xuất” trên màn hình chính</li> <li>2. Ứng dụng Unity xóa thông tin người dùng khỏi bộ nhớ đệm. Chuyển hướng người dùng đến trang đăng nhập.</li> </ol>	



Hình 2.3: Sơ đồ ca sử dụng chức năng quét hình ảnh mô hình

Bảng 2.4: Đặc tả sơ đồ ca chức năng quét hình ảnh hiển thị thông tin mô hình

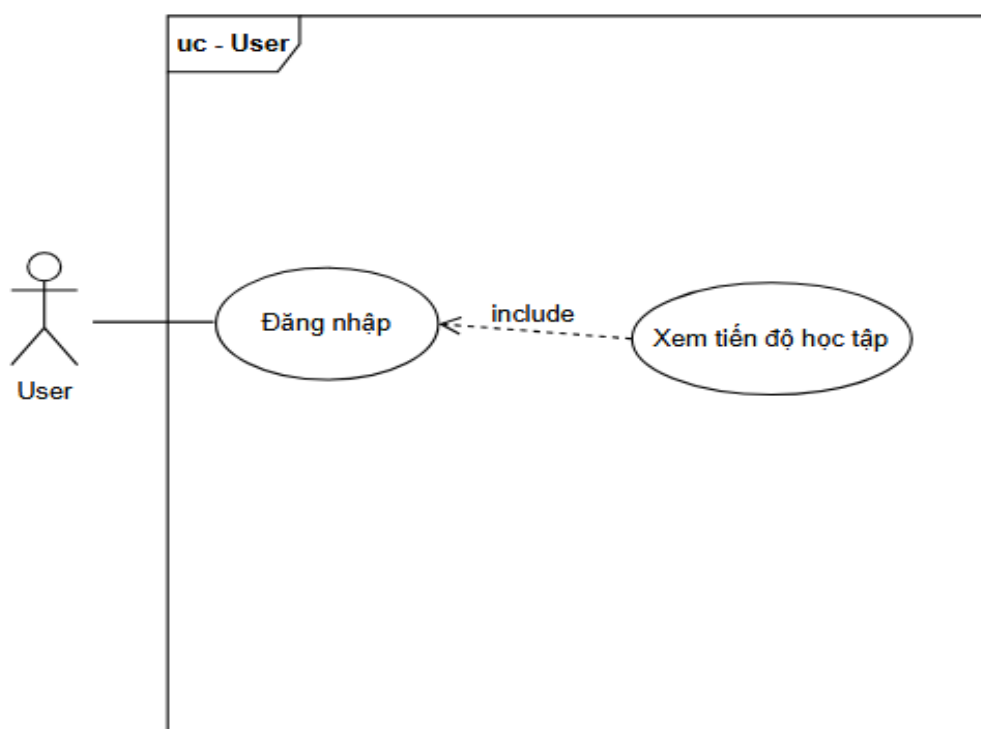
Tên Usecase	Người dùng xem thông tin mô hình thông qua việc quét hình ảnh mô hình
Tác nhân chính	User
Mô tả usecase	Người dùng xem mô hình, thông tin, câu chuyện liên quan tới mô hình, tương tác với mô hình và nghe phát âm của từ vựng.
Tiền điều kiện	Đã đăng nhập, đã chọn chủ đề từ vựng, đã quét hình ảnh thành công
Hậu điều kiện	Hiển thị mô hình, phát âm từ vựng của mô hình
Kích hoạt	Người dùng nhấn vào nút “Quét hình ảnh” trên trang chủ
Chuỗi sự kiện	<ol style="list-style-type: none"> <li>1. Người dùng nhấn vào nút “Quét mô hình” trên màn hình chính</li> <li>2. Chọn chủ đề từ vựng trong thanh tùy chọn chủ đề</li> <li>3. Quét hình ảnh thuộc chủ đề được chọn</li> <li>4. Hiển thị mô hình trên bề mặt ảnh được quét, phát âm từ vựng của mô hình</li> <li>5. Khi người dùng nhấn vào mô hình                         <ul style="list-style-type: none"> <li>- Phát âm từ vựng của mô hình</li> <li>- Hiển thị các bản trình chiếu thông tin (hình ảnh, văn bản) liên quan/ câu chuyện của mô hình</li> <li>- Nhấn nút “Vietnamese”/ “English” chuyển sang ngôn ngữ Việt Nam/ ngôn ngữ Anh, phát âm tiếng Việt/ tiếng Anh, chuyển đổi thông tin bản trình chiếu sang ngôn ngữ Việt/Anh</li> </ul> </li> <li>6. Khi người dùng kéo mô hình sang trái/phải thì mô hình xoay trái/phải</li> <li>7. Khi người dùng kéo mô to/ nhỏ, mô hình được phóng to ra/ thu bé lại</li> </ol>



Hình 2.4: Sơ đồ ca sử dụng chức năng luyện phát âm

Bảng 2.5: Đặc tả sơ đồ ca chức năng luyện phát âm kết hợp tương tác mô hình

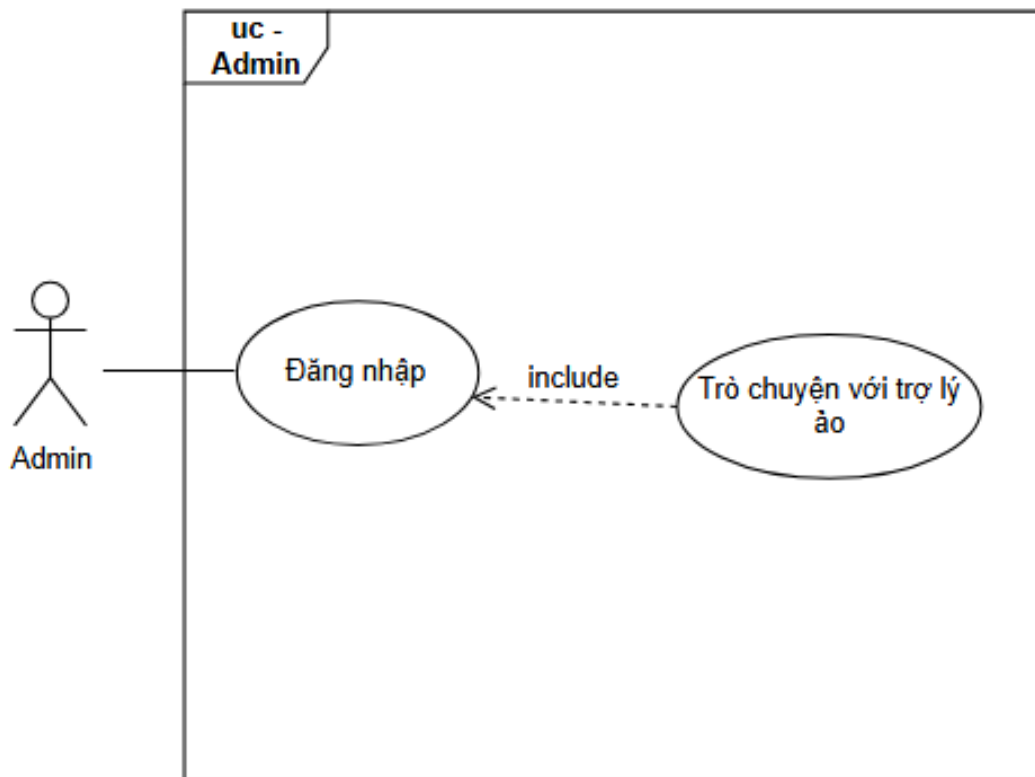
Tên Usecase	Người dùng luyện phát âm và tương tác với mô hình trong không gian thực
Tác nhân chính	User
Mô tả usecase	Người dùng nghe phát âm của mô hình, tương tác với mô hình và luyện phát âm từ vựng
Tiền điều kiện	Đã đăng nhập, đã quét bề mặt phẳng thành công, đã chọn chủ đề từ vựng
Hậu điều kiện	Hiện thị danh sách mô hình
Kích hoạt	Người dùng nhấn vào nút “Luyện phát âm” trên trang chủ
<p>Chuỗi sự kiện</p> <ol style="list-style-type: none"> <li>1) Người dùng nhấn vào nút “Luyện phát âm” trên trang chủ</li> <li>2) Chọn chủ đề từ vựng trong thanh tùy chọn chủ đề</li> <li>3) Chọn mô hình trong thanh tùy chọn mô hình</li> <li>4) Quét bề mặt phẳng xung quanh không gian thực</li> <li>5) Nhấn vào mặt phẳng để hiển thị mô hình trên không gian thực</li> <li>6) Nhấn vào nút “Vietnamese” để nghe phát âm tiếng Việt</li> <li>7) Nhấn vào nút “English” để nghe phát âm tiếng Anh</li> <li>8) Nhấn vào nút “Pronounce” để đọc và nhận kết quả phân tích được trả về từ server</li> <li>9) Nhấn vào nút “Disable” để tắt quét bề mặt</li> <li>10) Kết quả phát âm gồm từ nhận diện được, điểm số và kết quả được trả về hiển thị lên màn hình</li> <li>11) Khi người dùng kéo mô hình sang trái/phải thì mô hình xoay trái/phải</li> <li>12) Khi người dùng kéo mô to/ nhỏ, mô hình được phóng to ra/ thu bé lại</li> </ol>	



Hình 2.5: Sơ đồ ca sử dụng chức năng lưu trữ và theo dõi tiến độ học tập

Bảng 2.6: Đặc tả sơ đồ ca lưu trữ và xem tiến độ học tập

Tên Usecase	Người dùng xem tiến độ học tập cá nhân
Tác nhân chính	User
Mô tả usecase	Người dùng xem tiến độ học tập của cá nhân
Tiền điều kiện	Đã đăng nhập
Hậu điều kiện	Hiện thị tiến độ học tập theo từng mô hình
Kích hoạt	Người dùng nhấn vào nút “Tiến độ học tập” trên thanh điều hướng
Chuỗi sự kiện <ol style="list-style-type: none"> <li>1) Người dùng nhấn vào nút “Tiến độ học tập” trên màn hình chính</li> <li>2) Giao diện hiển thị tên các mô hình và tỉ lệ hoàn thành của từng mô hình</li> <li>3) Người dùng nhấn vào tên mô hình cụ thể</li> <li>4) Hiện thị tiến độ học chi tiết của mô hình bao gồm:                         <ul style="list-style-type: none"> <li>- Từ phát âm được nhận diện gần nhất</li> <li>- Điểm chấm phát âm</li> <li>- Tình trạng học, thời gian học các bản trình chiếu của mô hình</li> </ul> </li> </ol>	

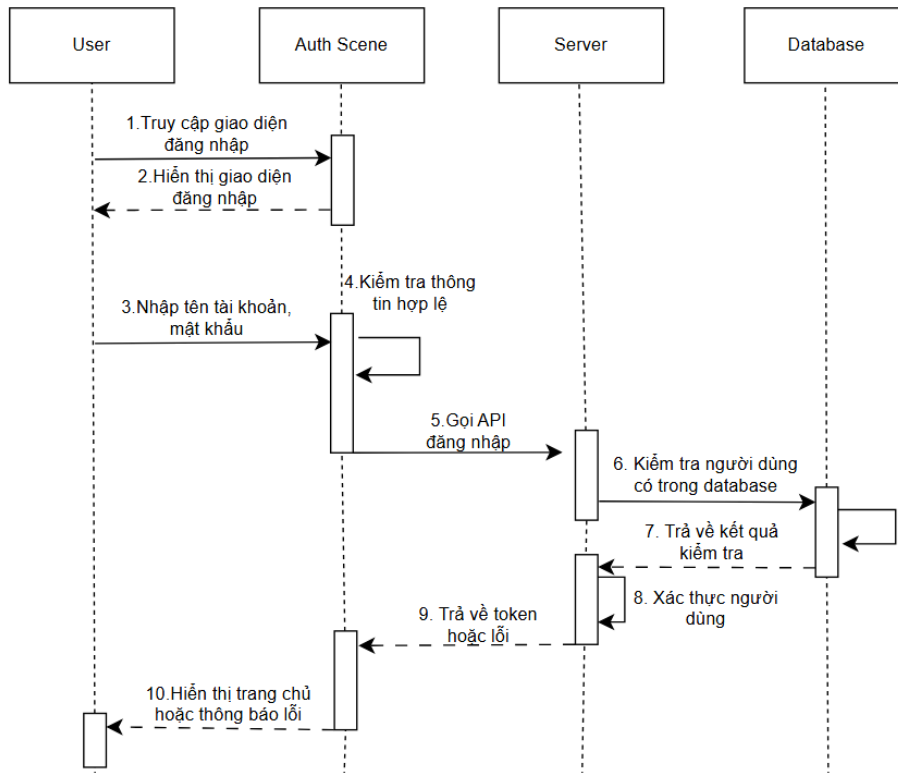


Hình 2.6: Sơ đồ ca sử dụng chức năng quản lý tài khoản

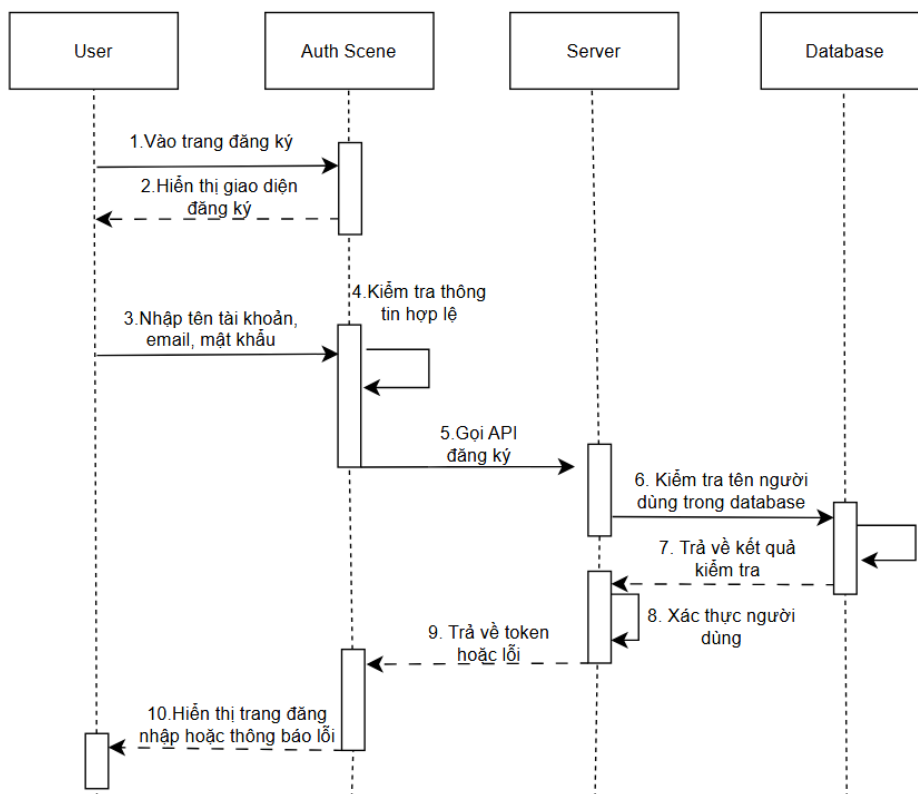
Bảng 2.7: Đặc tả sơ đồ ca chức năng quét trò chuyện với trợ lý ảo

Tên Usecase	Người dùng trò chuyện với trợ lý ảo
Tác nhân chính	User
Mô tả usecase	Người dùng trò chuyện với trợ lý ảo
Tiền điều kiện	Đã đăng nhập, chọn mục trợ lý ảo
Hậu điều kiện	Trò chuyện với trợ lý ảo
Kích hoạt	Người dùng nhấn vào nút “Trợ lý ảo” trên trang chủ
Chuỗi sự kiện	
<ol style="list-style-type: none"> <li>1) Nhấn vào nút “Trợ lý ảo”</li> <li>2) Gửi yêu cầu/ câu hỏi muốn hỏi</li> <li>3) Nhận kết quả trả về</li> </ol>	

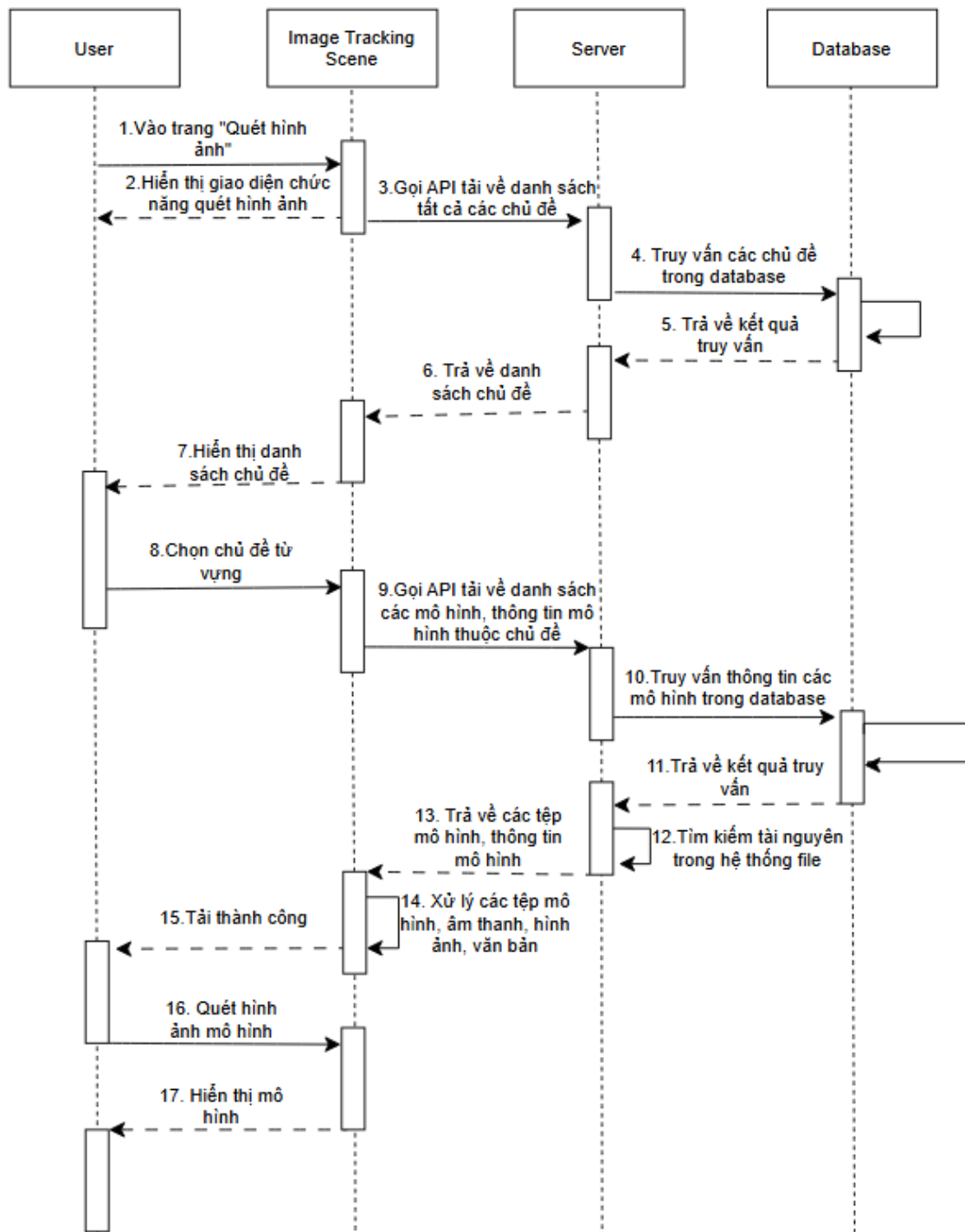
### 2.2.2. Sơ đồ tuần tự



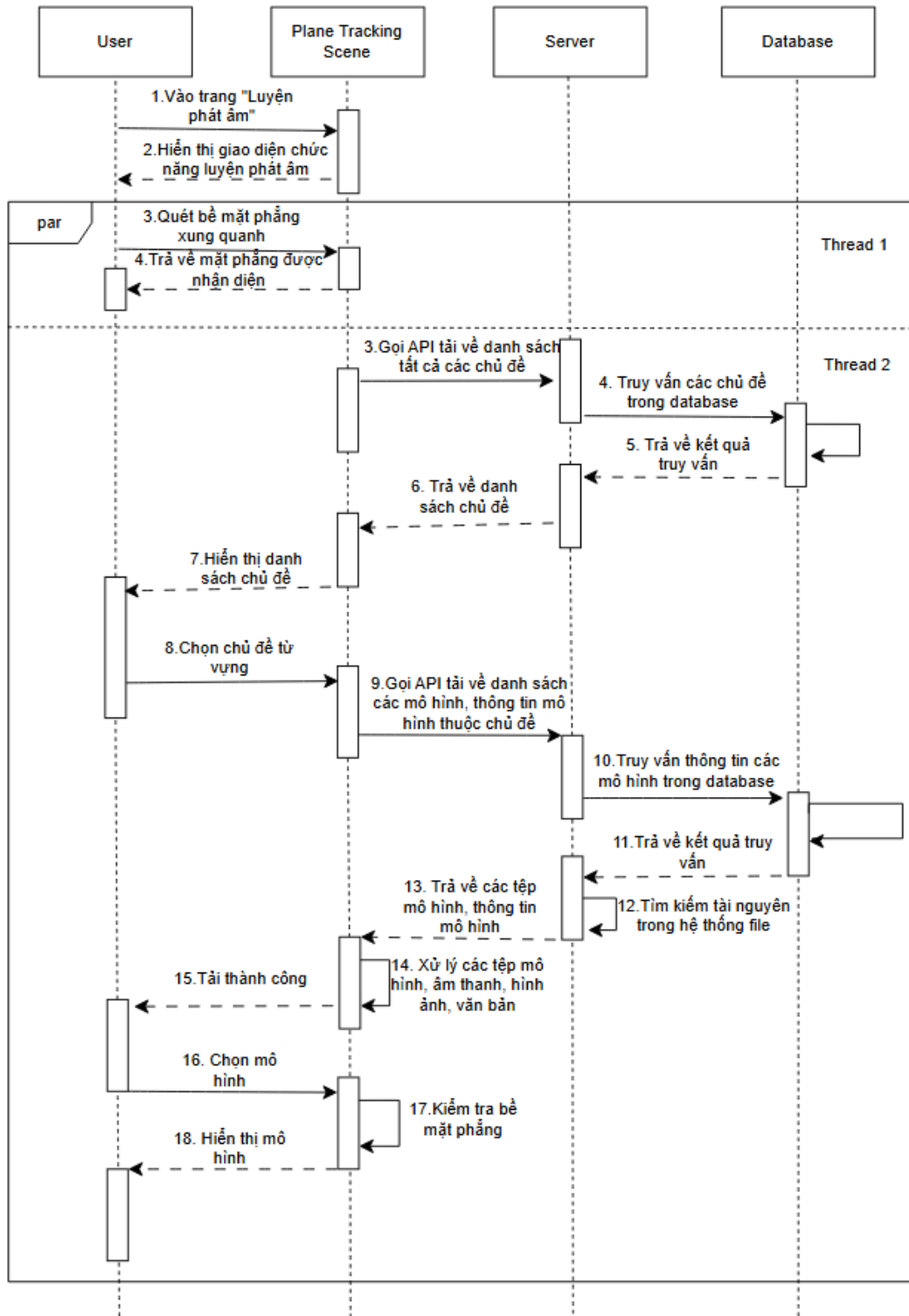
Hình 2.7: Sơ đồ tuần tự chức năng đăng nhập



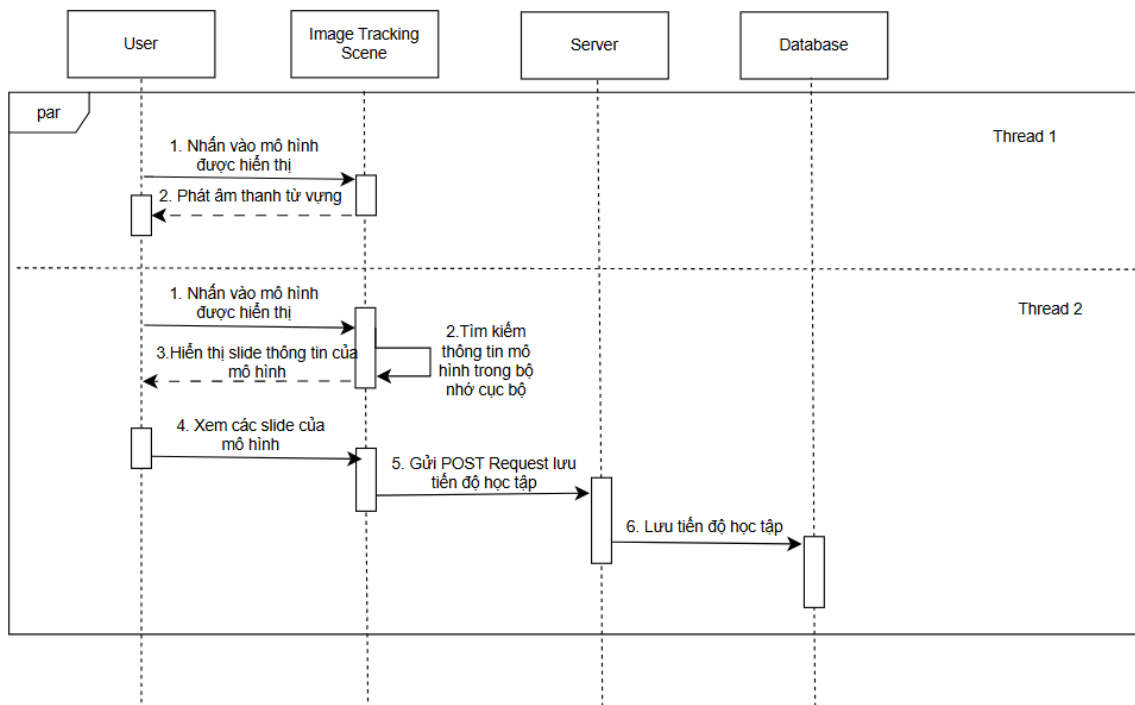
Hình 2.8: Sơ đồ tuần tự chức năng đăng ký



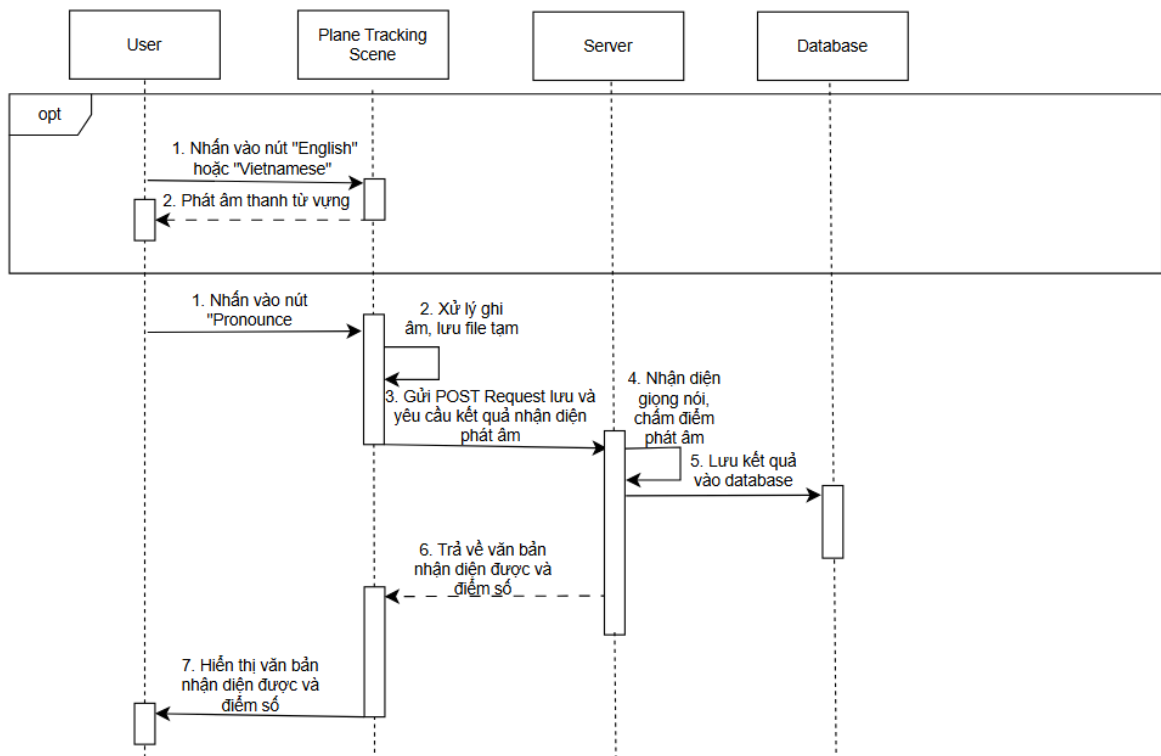
Hình 2.9: Sơ đồ tuần tự chức năng “Thông tin mô hình”



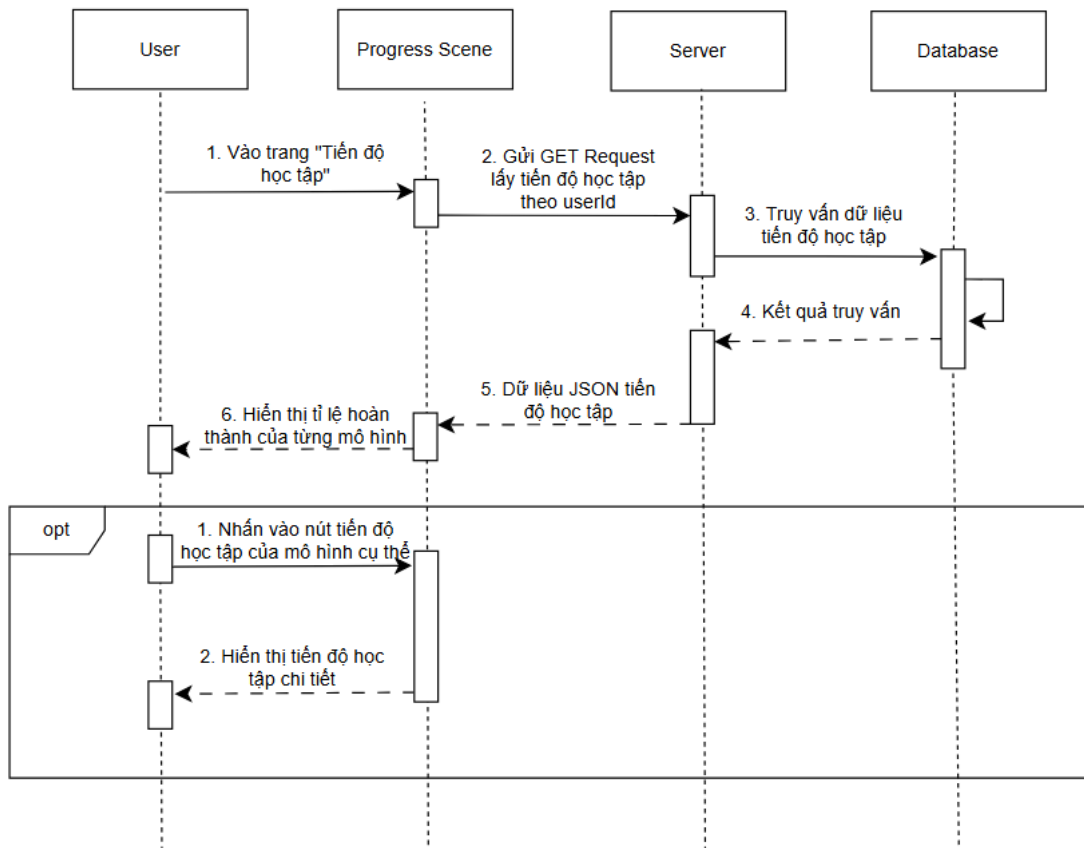
Hình 2.10: Sơ đồ tuần tự chức năng “Luyện phát âm”



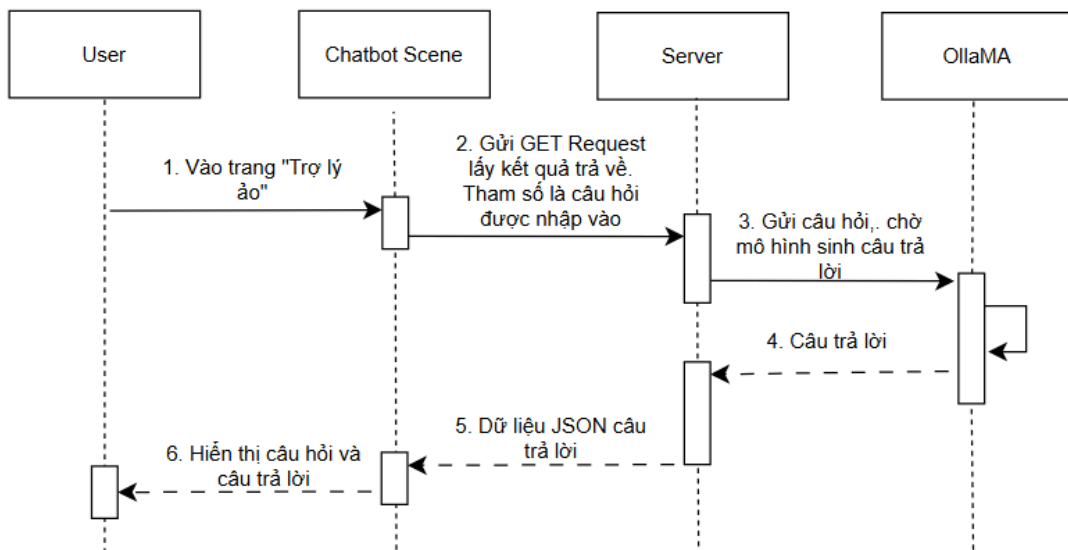
Hình 2.11: Sơ đồ tuần tự chức năng hiển thị bản trình chiếu thông tin mô hình



Hình 2.12: Sơ đồ tuần tự chức năng chấm điểm phát âm

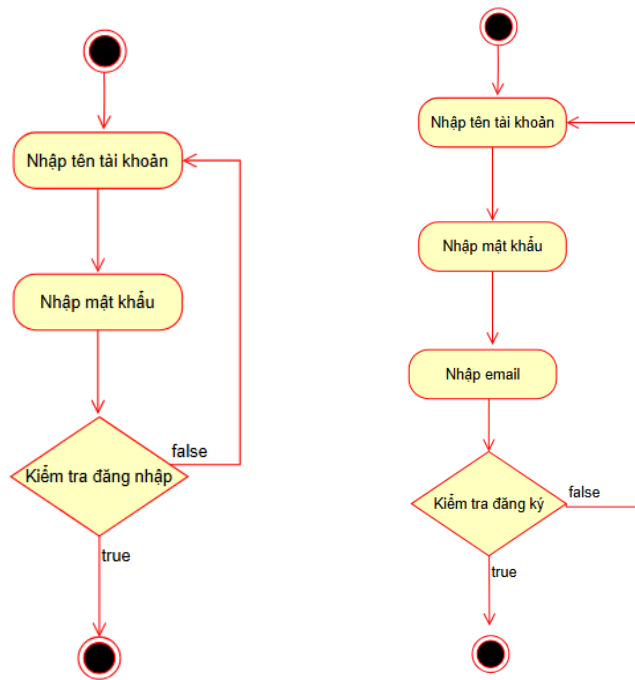


Hình 2.13: Sơ đồ tuần tự chức năng “Tiến độ học tập”

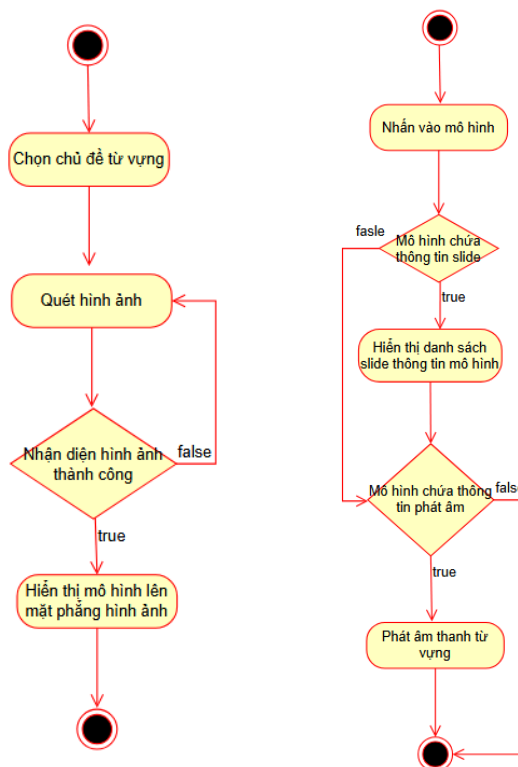


Hình 2.14: Sơ đồ tuần tự chức năng “Trợ lý ảo”

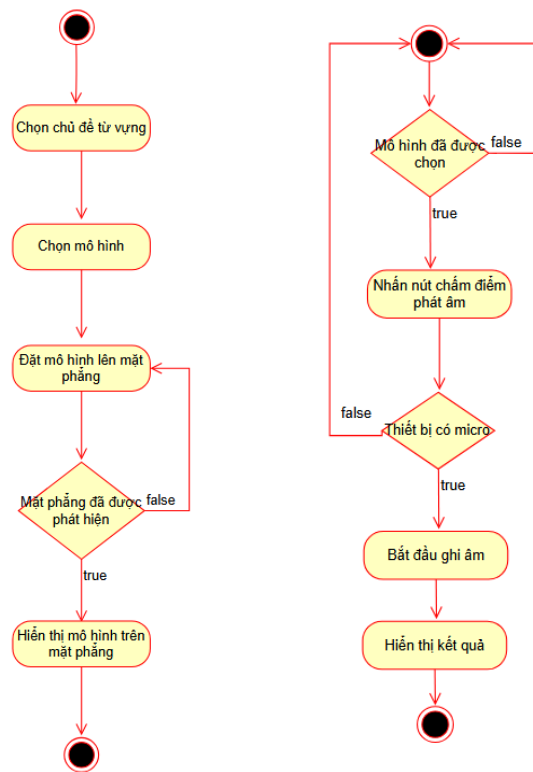
### 2.2.3. Sơ đồ nguyên lý hoạt động



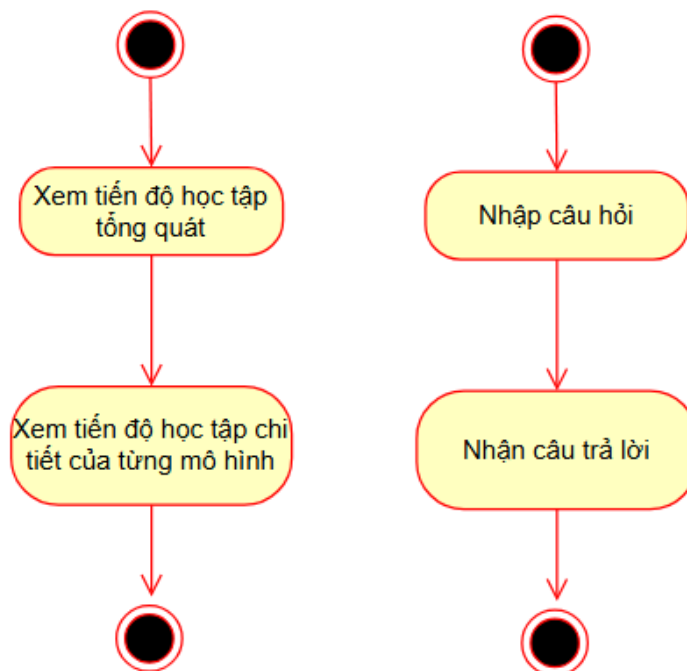
Hình 2.15: Sơ đồ hoạt động chức năng đăng nhập/ đăng ký



Hình 2.16: Sơ đồ hoạt động chức năng “Thông tin mô hình”



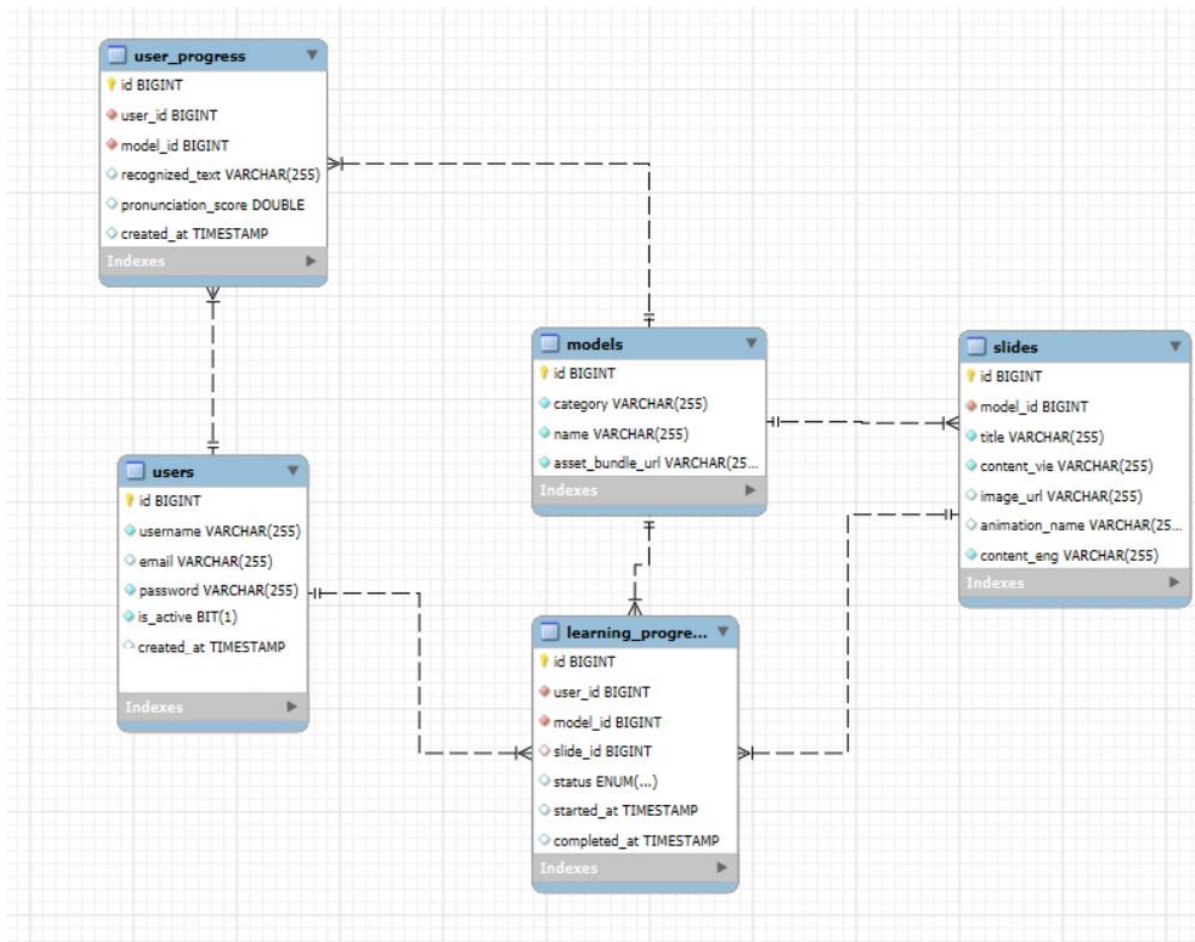
Hình 2.17: Sơ đồ hoạt động chức năng “Luyện phát âm”



Hình 2.18: Sơ đồ hoạt động chức năng “Tiến độ học tập” và “Trợ lý ảo”

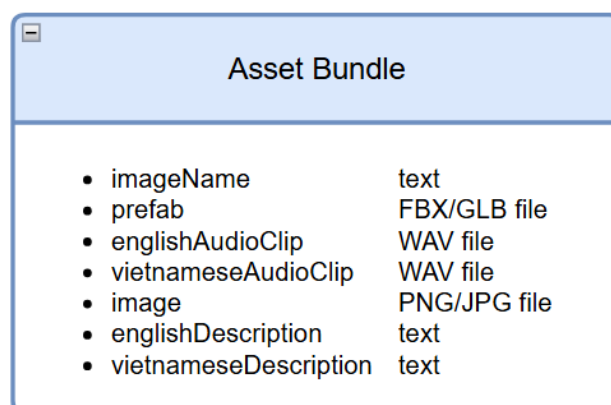
## 2.3. Thiết kế cơ sở dữ liệu

### 2.3.1. Cơ sở dữ liệu



Hình 2.19: Sơ đồ thiết kế cơ sở dữ liệu

Với tập Asset Bundle gồm các dữ liệu sau:



Hình 2.20: Mô tả các dữ liệu có trong các tập Asset Bundle

### 2.3.2. Chi tiết các bảng trong cơ sở dữ liệu

Bảng users: Lưu trữ thông tin tài khoản người dùng

Bảng 2.8: Bảng users

Cột	Kiểu dữ liệu	Ràng buộc	Mô tả
id	BIGINT	PRIMARY KEY	ID duy nhất của người dùng.
username	VARCHAR(255)	NOT NULL, UNIQUE	Tên tài khoản người dùng
email	VARCHAR(255)	NOT NULL	Email người dùng
password	VARCHAR(255)	NOT NULL	Mật khẩu tài khoản
is_active	BIT(1)	NOT NULL	Trạng thái hoạt động
created_at	TIMESTAMP	NOT NULL	Thời gian tạo tài khoản

Bảng models: Lưu trữ thông tin các mô hình 3D

Bảng 2.9: Bảng models

Cột	Kiểu dữ liệu	Ràng buộc	Mô tả
id	BIGINT	PRIMARY KEY	ID duy nhất của mô hình 3D
category	VARCHAR(255)	NOT NULL	Danh mục từ vựng
name	VARCHAR(255)	NOT NULL	Tên mô hình 3D
asset_bundle_url	VARCHAR(255)	NOT NULL	Đường dẫn tới tệp asset bundle chứa thông tin mô hình trên server

Bảng slides: Lưu trữ thông tin các bản trình chiếu của từng mô hình 3D

Bảng 2.10: Bảng slides

Cột	Kiểu dữ liệu	Ràng buộc	Mô tả
-----	--------------	-----------	-------

id	BIGINT	PRIMARY KEY	ID duy nhất của bản thuyết trình
model_id	BIGINT	NOT NULL, UNIQUE, FOREIGN KEY	ID mô hình 3D
title	VARCHAR(255)	NOT NULL	Tiêu đề của bản thuyết trình
content_vie	VARCHAR(255)	NOT NULL	Thông tin của bản thuyết trình – tiếng Việt
content_eng	VARCHAR(255)	NOT NULL	Thông tin của bản thuyết trình – tiếng Anh
image_url	VARCHAR(255)	NOT NULL	Đường dẫn tới tệp hình ảnh mô tả bản thuyết trình trên server
animation_name	VARCHAR(255)	NOT NULL	Tên hoạt ảnh mô hình sử dụng

Bảng user\_progress: Lưu trữ tiến độ luyện phát âm của người dùng

*Bảng 2.11: Bảng user\_progress*

Cột	Kiểu dữ liệu	Ràng buộc	Mô tả
id	BIGINT	PRIMARY KEY	ID duy nhất của bản ghi tiến độ học phát âm
user_id	BIGINT	NOT NULL, UNIQUE, FOREIGN KEY	ID của người dùng
model_id	BIGINT	NOT NULL, UNIQUE, FOREIGN KEY	ID mô hình 3D
recognized_text	VARCHAR(255)	NOT NULL	Văn bản nhận diện được từ bản ghi âm của người dùng

pronunciation_score	DOUBLE	NOT NULL	Điểm phát âm
created_at	TIMESTAMP	NOT NULL	Thời gian tạo bản ghi

Bảng learning\_progress: Lưu trữ tiến độ học các bản thuyết trình của người dùng

*Bảng 2.12: Bảng learning\_progress*

Cột	Kiểu dữ liệu	Ràng buộc	Mô tả
id	BIGINT	PRIMARY KEY	ID duy nhất của bản ghi tiến độ học các bản thuyết trình
user_id	BIGINT	NOT NULL, UNIQUE, FOREIGN KEY	ID của người dùng
model_id	BIGINT	NOT NULL, UNIQUE, FOREIGN KEY	ID mô hình 3D
slide_id	BIGINT	NOT NULL, UNIQUE, FOREIGN KEY	ID bản thuyết trình
status	ENUM ('STARTED', 'COMPLETED')		Tiến độ học tập, gồm 2 trạng thái là “Đã bắt đầu” và “Đã hoàn thành”
started_at	TIMESTAMP	NOT NULL	Thời gian bắt đầu học
completed_at	TIMESTAMP		Thời gian hoàn thành bài học

## 2.4. Kết chương

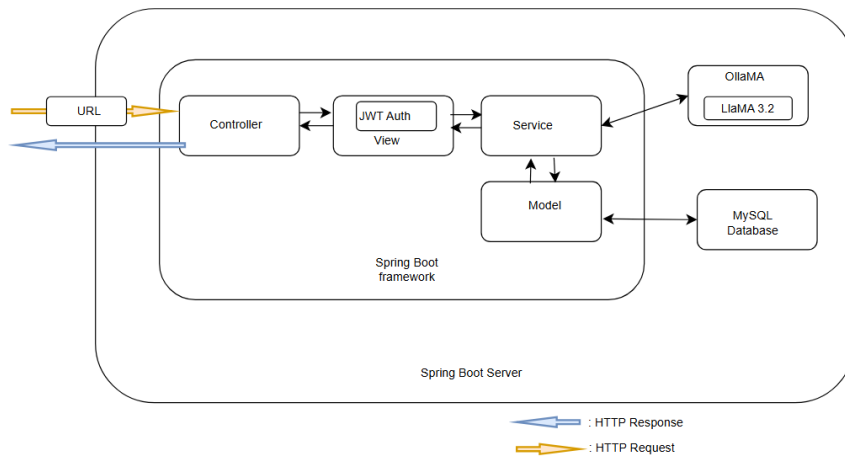
Chương này trình bày các sơ đồ ca sử dụng, sơ đồ tuần tự, sơ đồ nguyên lý hoạt động của các chức năng trong chương trình. Các sơ đồ đã mô tả các chuỗi hành động cũng như cách thức, thứ tự tương tác giữa người dùng với ứng dụng di động và từ ứng dụng di động tới server. Kế đó là bản thiết kế cơ sở dữ liệu, các bảng mô tả các thành phần trong cơ sở dữ liệu. Qua đó, luồng hoạt động và cách thức xử lý dữ liệu đã được trình bày rõ ràng và đầy đủ.

## CHƯƠNG 3: TRIỂN KHAI VÀ KẾT QUẢ ĐẠT ĐƯỢC

### 3.1. Server

Công nghệ sử dụng: Spring Boot, Java, Maven, MySQL, REST API

Môi trường phát triển: Eclipse



Hình 3.1: Thiết kế kiến trúc server Spring Boot

Sử dụng Spring Boot framework để xây dựng các API RESTful hỗ trợ giao tiếp giữa máy chủ và ứng dụng di động. Dùng JWT để xác thực người dùng. Chuẩn dữ liệu để trao đổi dữ liệu người dùng là JSON.

#### 3.1.1. Cài đặt các lớp truy xuất dữ liệu

Model, Entity và Repository là những lớp dùng để thực hiện truy xuất và chuyển đổi dữ liệu thành dữ liệu phù hợp với ứng dụng, chứa các trường chứa thông tin. Tương ứng với các bảng `user_progress`, `models`, `slides`, `learning_progress`, `users` trong cơ sở dữ liệu là các lớp ánh xạ và chuyển đổi dữ liệu được cài đặt trên server được Spring Boot framework quản lý.

```
@Table(name = "models")
public class Model {

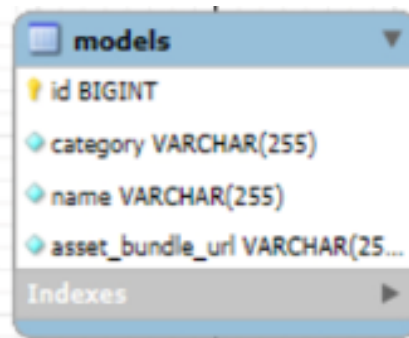
    @Id
    @GeneratedValue(strategy = GenerationType.IDENTITY)
    private Long id;

    @Column(nullable = false)
    private String category;

    @Column(nullable = false)
    private String name;

    @Column(name = "asset_bundle_url", nullable = false)
    private String assetBundleUrl;

    @OneToMany(mappedBy = "model", cascade = CascadeType
    private List<Slide> slides;
}
```



Hình 3.2: Minh họa bảng models trong cơ sở dữ liệu và lớp Model ánh xạ

Danh sách các lớp ánh xạ được cài đặt:

*Bảng 3.1: Các lớp ánh xạ được sử dụng để truy xuất thông tin từ cơ sở dữ liệu*

Tên repository	Chức năng
LearningProgressRepository	Lưu trữ - truy xuất thông tin tiến độ học tập người dùng
ModelRepository	Lưu trữ - truy xuất thông tin các mô hình
SlideRepository	Lưu trữ - truy xuất thông tin bản trình chiếu của từng mô hình
UserProgressRepository	Lưu trữ - truy xuất thông tin học phát âm người dùng
UserRepository	Lưu trữ - truy xuất thông tin người dùng

### 3.1.2. Cài đặt các lớp xử lý logic

Service là tập hợp các class dùng để xử lý logic khi truyền dữ liệu giữa server và ứng dụng di động. Trong hệ thống được xây dựng, các lớp này dùng để chuyển đổi dữ liệu giữa json hoặc multipart-form thành các kiểu dữ liệu mà Spring Boot có thể quản lý và ngược lại, thêm điều kiện lọc thông tin khi truyền nhận dữ liệu và xử lý, xử lý các tệp trong hệ thống tập tin của server. Ngoài ra còn là lớp truy cập đến các service ngoài Spring Boot như Faster Whisper hay mô hình ngôn ngữ LLaMA

*Bảng 3.2: Các lớp xử lý logic của hệ thống*

Tên service	Chức năng
AuthService	Xử lý logic đăng nhập - đăng ký cho hệ thống
CustomUserDetailsService	Lớp triển khai giao diện UserDetailsService của Spring Security, hỗ trợ cho lớp AuthService trong việc xử lý đăng nhập, đăng ký cho hệ thống
FileStorageService	Xử lý logic nhận tệp ghi âm phát âm của người dùng
LearningProgressService	Xử lý logic lưu, truy xuất thông tin tiến độ học tập của người dùng
ModelService	Xử lý logic truy xuất danh sách mô hình
SpeechRecognitionService	Xử lý logic chấm điểm phát âm, gọi dịch vụ Faster Whisper

### 3.1.3. Cài đặt các lớp routing

Các lớp thuộc gói Controller là các lớp xử lý việc routing cho hệ thống. Để chuyển hướng các request đến các lớp Service cần thiết, cần phải cung cấp cho các lớp

Controller tên đường dẫn URL tương ứng với từng request, kèm theo đó là các tham số cần thiết để Service hoạt động như ý muốn

Bảng 3.3: Danh sách các URL được lớp Controller quản lý

URL	Phương thức HTTP	Tham số	Chức năng
/api/auth/register	POST	Lớp AuthRequest (username, password,email, role) thuộc gói Model	Chuyển hướng đến dịch vụ đăng ký
/api/auth/login	POST	Lớp AuthRequest (username, password) thuộc gói Model	Chuyển hướng đến dịch vụ đăng nhập
/api/chatbot	GET	Tin nhắn của user	Trả về câu trả lời của mô hình LLaMa
/api/speech/recognize	POST	Tệp ghi âm phát âm (tệp WAV) Từ vựng chính xác Mã người dùng Mã mô hình	Chuyển hướng đến dịch vụ chuyển giọng nói thành văn bản và chấm điểm phát âm
/api/progress	POST	Mã bản trình chiếu Tình trạng học Mã người dùng Mã mô hình	Chuyển hướng đến dịch vụ lưu trữ tiến độ học tập người dùng
/api/progress/get/{userId}	GET	Mã người dùng	Chuyển hướng đến dịch vụ truy xuất tiến độ học của người dùng
/api/models/categories	GET		Chuyển hướng đến dịch vụ truy

			xuất danh sách danh mục từ vựng
/api/models/category/{categories}	GET	Tên danh mục	Chuyển hướng đến dịch vụ truy xuất các mô hình thuộc danh mục được chọn

Các request gửi tới các endpoint có thể kèm theo payload là những dữ liệu định dạng json, form-multipart với auth là bearer token

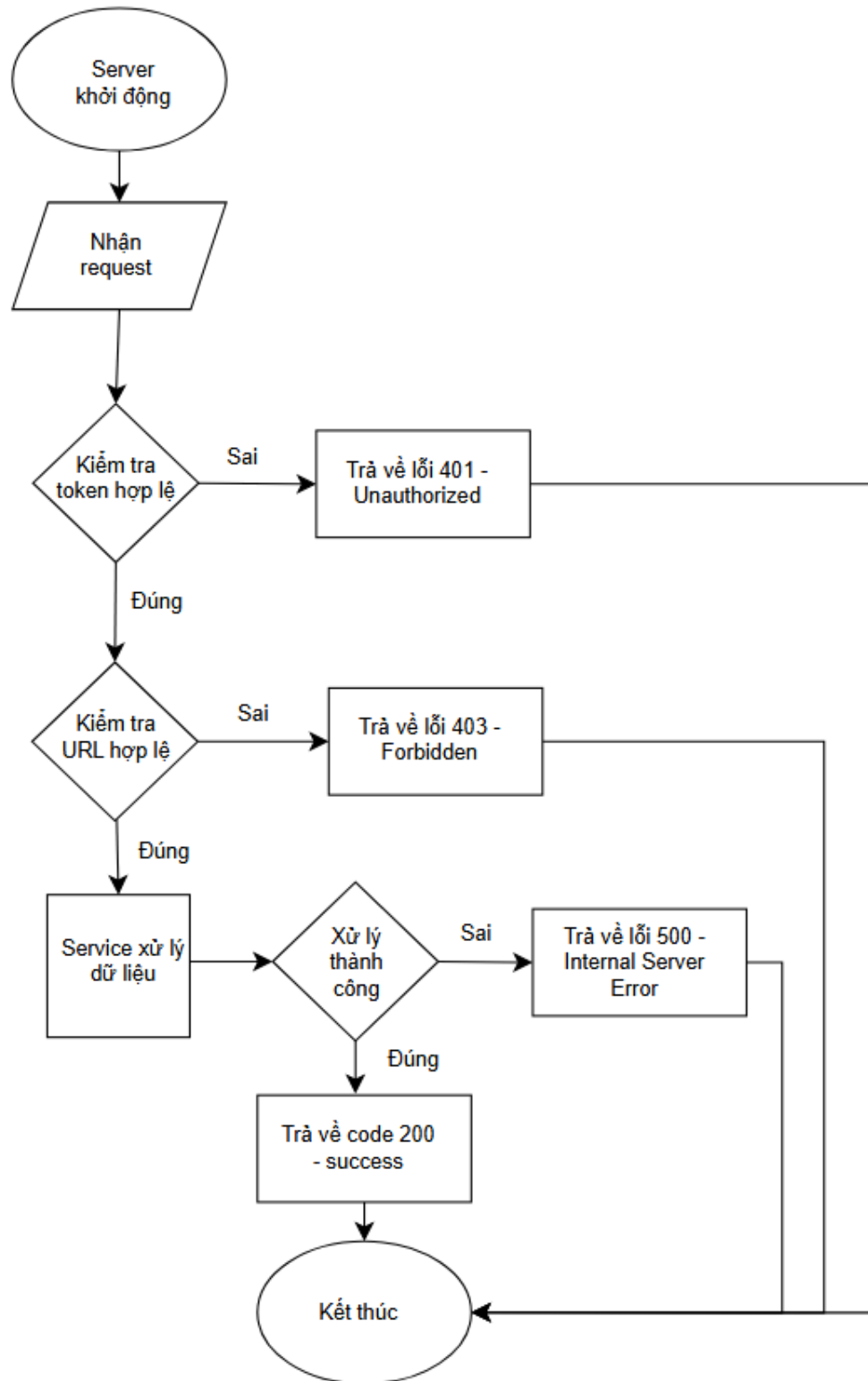
Ví dụ để tạo thông tin một người dùng mới thì các thông tin kèm theo của request bao gồm:

- URL: [host-address:port]/api/auth/register /
- HTTP method: POST
- Content-type: JSON
- Payload: {"username": "testuser7", "password": "password123", "email": "trungaaa@gmail.com", "role" : 1}

#### **3.1.4. Xác thực người dùng sử dụng JWT**

Sử dụng thư viện JsonWebToken kết hợp Spring Security để tạo dịch vụ xác thực người dùng. Khi đăng nhập hay đăng ký thành công thì server sẽ trả về một token hỗ trợ xác thực không cần tài khoản. Thời gian tồn tại của token là 24 giờ. Khi token hết hạn thì người dùng phải đăng nhập lại để có thể truy cập đến các chức năng của mô hình

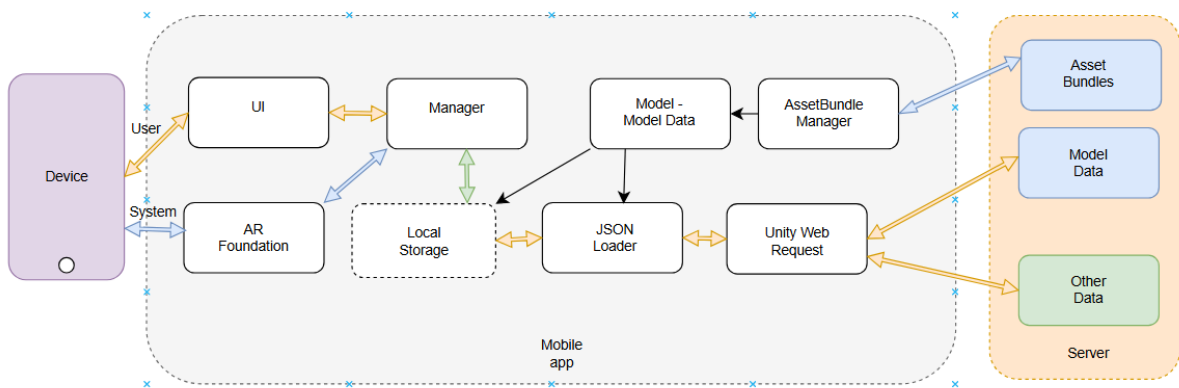
#### **3.1.5. Sơ đồ xử lý các request**



Hình 3.3: Sơ đồ xử lý request của server

### 3.2. Ứng dụng di động

- Phiên bản AR Foundation sử dụng: 6.0.5
- Phiên bản Unity sử dụng: Unity 6 (6000.0.44f1)



Hình 3.4: Mô tả kiến trúc ứng dụng di động

Ứng dụng di động bao gồm:

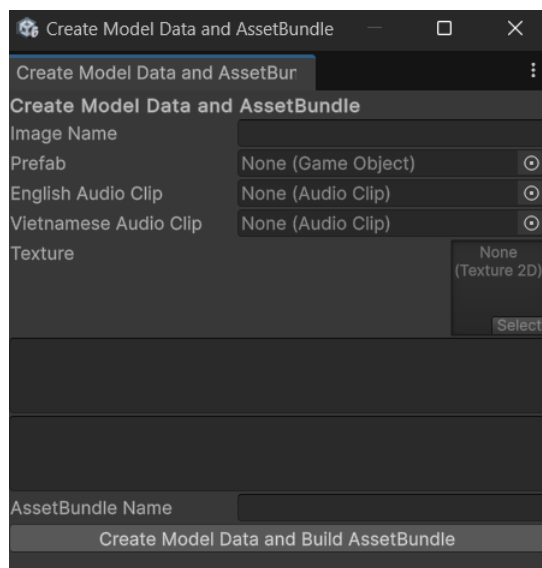
- UI sẽ tương tác trực tiếp với người dùng
- API AR Foundation sẽ trao đổi với phía hệ thống thông qua camera, các cảm biến tích hợp trong thiết bị di động
- Manager là thành phần dùng để xử lý các yêu cầu của người dùng thông qua giao diện và giao tiếp với các API AR Foundation khi người dùng tương tác với mô hình trong không gian thực.

Về phần giao tiếp với server, sử dụng:

- Thư viện Unity Networking với AssetBundle để tải tệp tài liệu chứa mô hình được đóng gói bởi Unity Editor và thông tin mô hình, Unity Web Request để truyền nhận dữ liệu liên quan tới tiến trình học, thông tin bản trình chiếu, thông tin mô hình, thông tin người dùng.
- JSON Loader thuộc thư viện Newtonsoft.Json để chuyển hóa dữ liệu JSON thành dữ liệu phù hợp cho hệ thống

### 3.2.1. Lưu trữ mô hình và thông tin mô hình

Sử dụng Asset Bundle để lưu trữ mô hình và thông tin mô hình

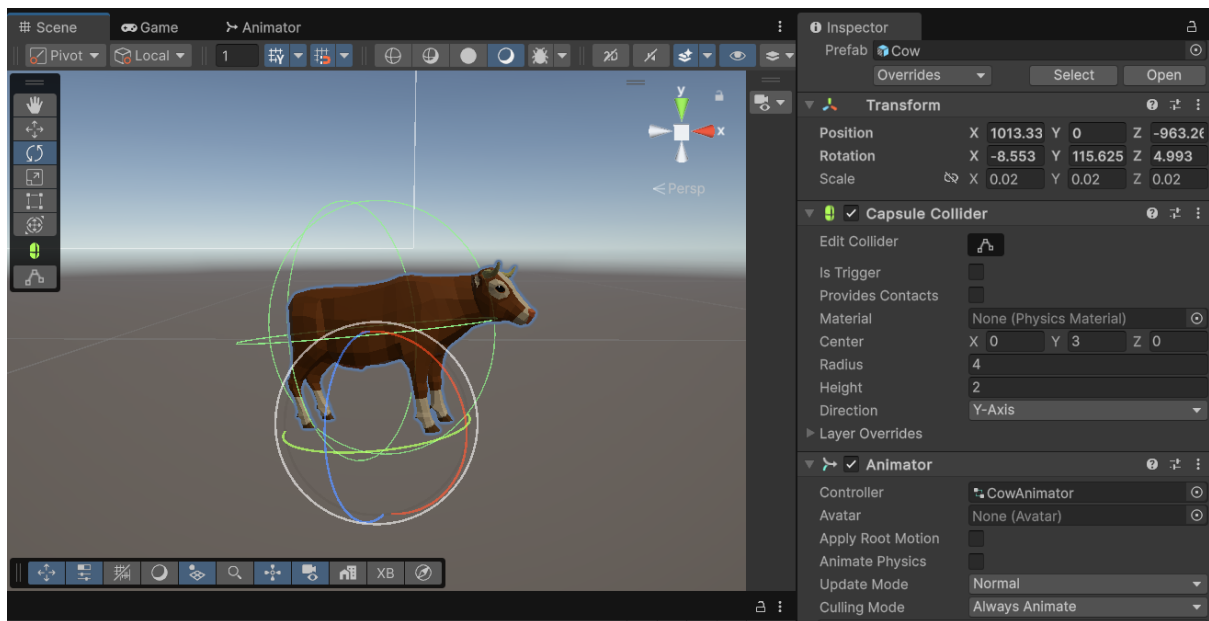


Hình 3.5: Cửa sổ tạo AssetBundle với mô hình và thông tin mô hình

Tạo một script kế thừa lớp EditorWindow để tạo một cửa sổ tạo tệp AssetBundle gồm mô hình và thông tin mô hình. Điều này giúp cho mô hình có thể xuất ra thành tệp có thể truyền tải giữa client và server, hỗ trợ xử lý lúc runtime. Thông tin trong AssetBundle bao gồm

- Image Name: Tên ảnh nhận diện của ảnh chứa thông tin mô hình trong ARReferenceLibraryImage
- Prefab: Mẫu mô hình, đã được điều chỉnh độ lớn, thêm collider để người dùng có thể tương tác, vị trí và góc độ được điều chỉnh hợp lý thông qua Unity Inspector.
- English audio clip: Tệp âm thanh phát âm tiếng Anh chuẩn của mô hình
- Vietnamese audio clip: Tệp âm thanh phát âm tiếng Việt chuẩn của mô hình
- Texture: Ảnh 2D của mô hình.
- English description: Mô tả tiếng Anh của mô hình
- Vietnamese description: Mô tả tiếng Việt của mô hình
- AssetBundle name: Tên AssetBundle

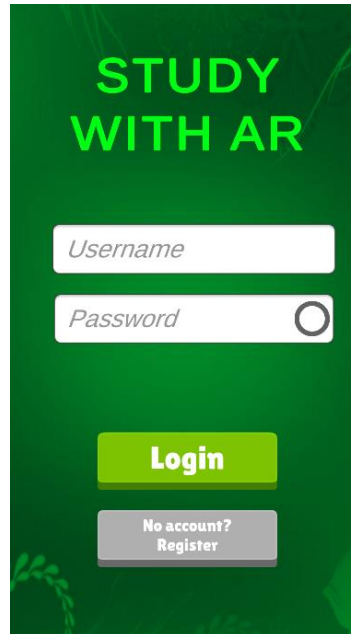
Sau khi tạo thành công AssetBundle, tệp sẽ được lưu vào hệ thống tập tin của server và truyền về ứng dụng di động khi có request



Hình 3.6: Cửa sổ để điều chỉnh mô hình gốc thành prefab cho AssetBundle

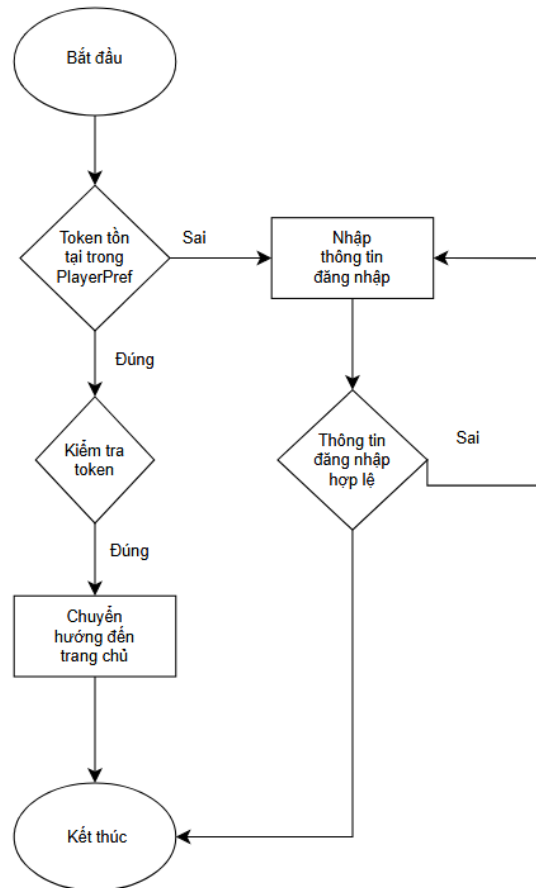
### 3.2.2. Đăng nhập

- Màn hình đăng nhập gồm có trường nhập tên tài khoản, mật khẩu, tùy chọn hiển thị mật khẩu, nút đăng nhập và tùy chọn sang màn hình đăng ký.



Hình 3.7: Biểu mẫu đăng nhập của ứng dụng

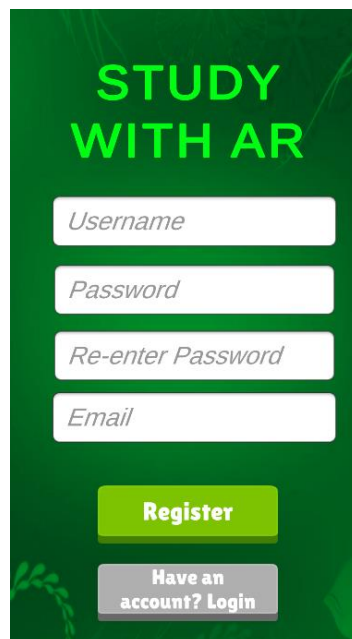
- Khi người dùng truy cập ứng dụng lần đầu, ứng dụng sẽ hiển thị biểu mẫu đăng nhập. Tính năng này sẽ sử dụng JWT token, giúp cho server không cần phải lưu trạng thái người dùng và thay cho session và cookies để ghi nhớ phiên đăng nhập và xác thực các yêu cầu xử lý dữ liệu.
- Nếu người dùng đã đăng nhập và token vẫn được lưu trong danh mục PlayerPrefs, ứng dụng sẽ tự động xác thực đăng nhập bằng token mà không cần đăng nhập lại.
- Sơ đồ thể hiện quá trình xử lý đăng nhập được mô tả như sau:



Hình 3.8: Sơ đồ xử lý đăng nhập

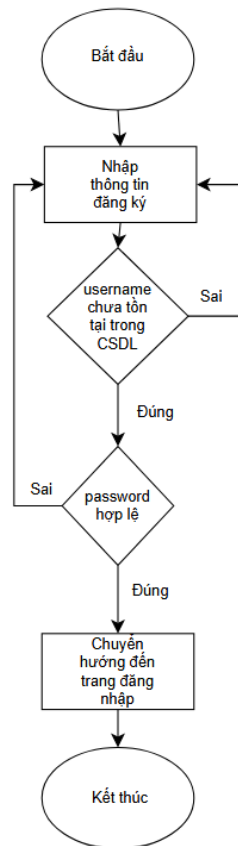
### 3.2.3. Đăng ký

Màn hình đăng ký gồm có trường nhập tên tài khoản, mật khẩu, email nút đăng ký và tùy chọn quay về đăng nhập.



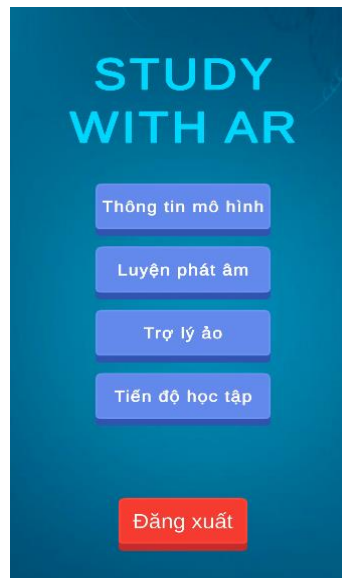
Hình 3.9: Biểu mẫu đăng ký của ứng dụng

Khi người dùng chuyển hướng sang trang ký, ứng dụng sẽ hiển thị biểu mẫu đăng ký để người dùng tạo tài khoản mới



Hình 3.10: Sơ đồ xử lý đăng ký

### 3.2.4. Màn hình chính

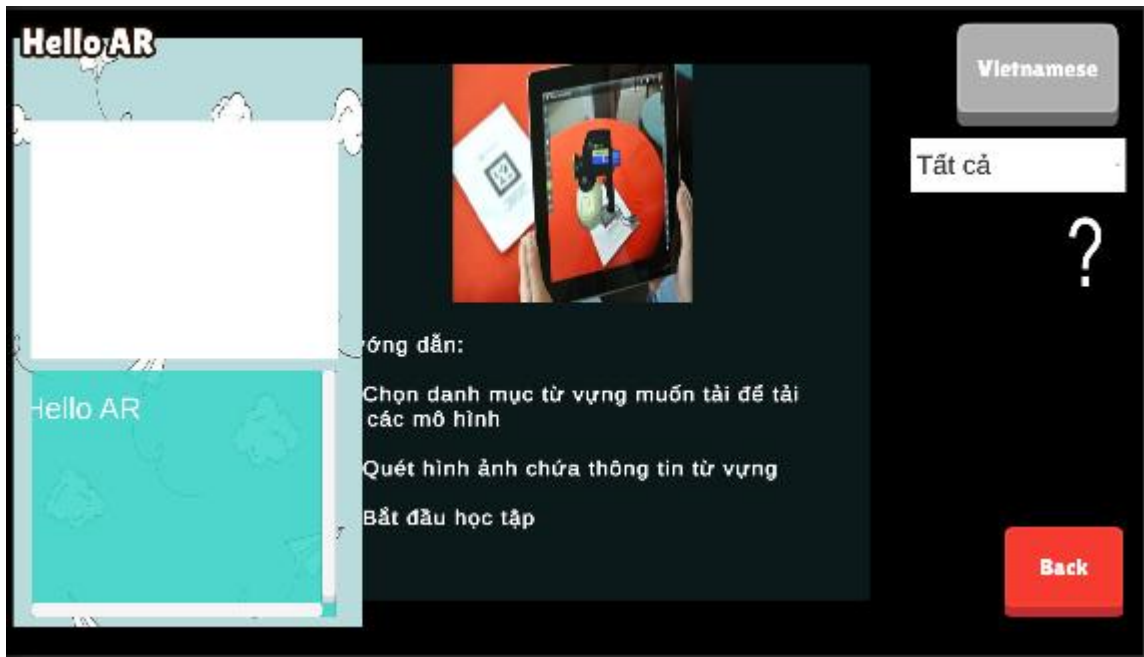


Hình 3.11: Giao diện màn hình trang chủ

Màn hình trang chủ gồm một giao diện điều hướng đến từng chức năng của ứng dụng và nút đăng xuất.

### 3.2.5. Chức năng “Thông tin mô hình”

Màn hình chức năng xem thông tin mô hình mặc định gồm danh sách tùy chọn danh mục mô hình, nút hướng dẫn sử dụng chức năng xem thông tin và nút quay lại.



Hình 3.12: Scene gốc chức năng “Thông tin mô hình”

Khi người dùng chọn danh mục từ vừng muốn xem thông tin, phân cảnh sẽ chịu trách nhiệm tải về mô hình và thông tin các mô hình đó để truyền sang cho thư viện AR Foundation xử lý

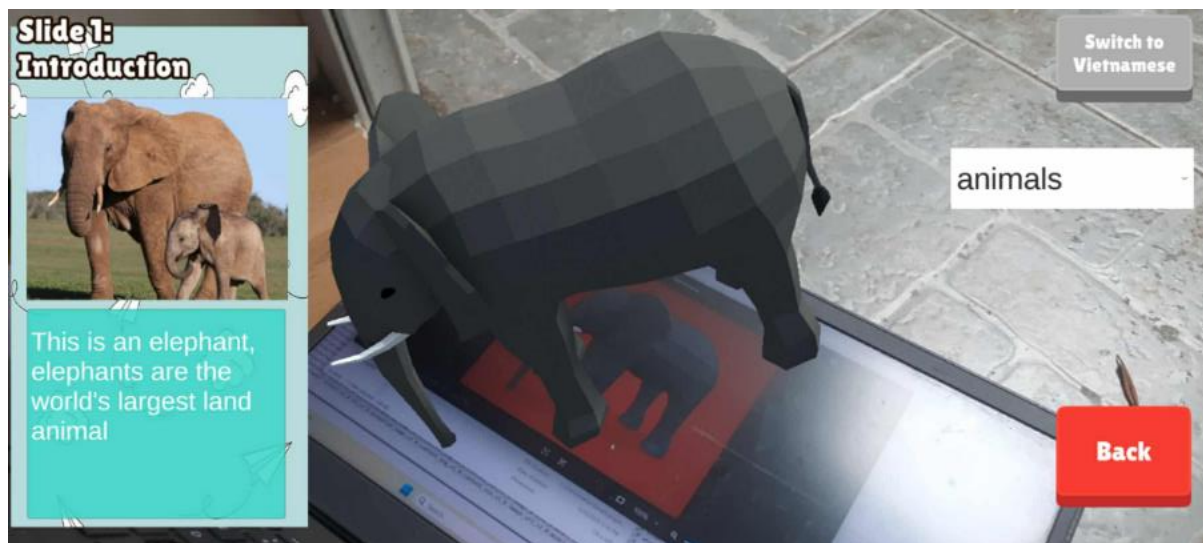


Hình 3.13: Màn hình chờ tải mô hình sau khi chọn danh mục “animals”

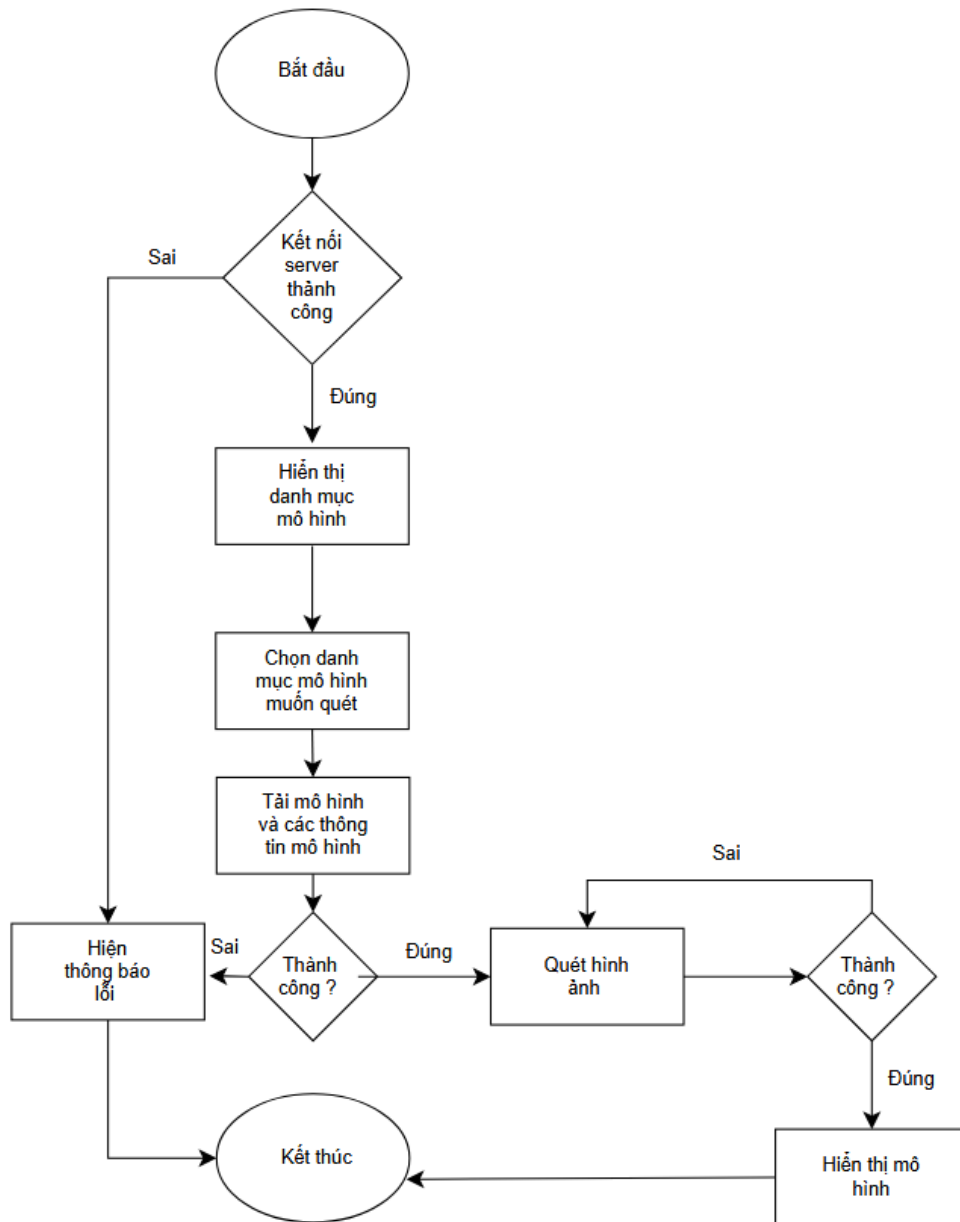
Trong trường hợp người dùng quét thành công mô hình từ hình ảnh và tương tác với mô hình, giao diện chức năng sẽ hiển thị thêm giao diện phụ để trình bày thông tin mô hình và nút chuyển đổi ngôn ngữ.



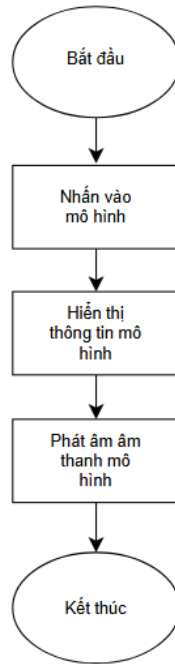
*Hình 3.14: Minh họa hiển thị mô hình trên hình ảnh được quét sau khi tải mô hình và quét hình ảnh thành công*



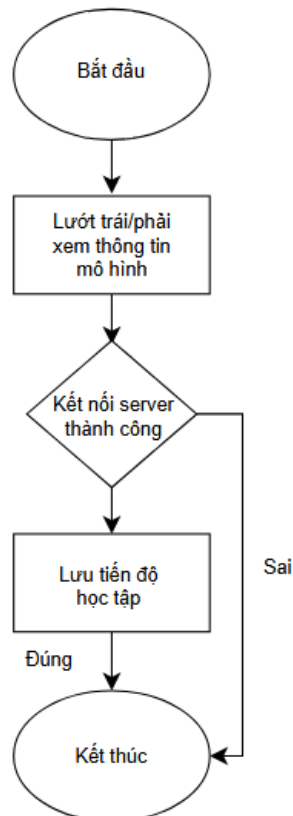
*Hình 3.15: Hiển thị thông tin mô hình khi nhấn vào mô hình*



Hình 3.16: Sơ đồ xử lý tải mô hình và hiển thị mô hình



Hình 3.17: Sơ đồ xử lý hiển thị thông tin và phát âm thanh của mô hình



Hình 3.18: Sơ đồ xử lý hiển thị thông tin và phát âm thanh của mô hình

Chức năng lưu tiến độ học tập là chức năng hoạt động ngầm, nó chỉ được gọi khi người dùng có hành động lướt trái phải để xem thông tin mô hình. Thông tin được lưu lại bao gồm tình trạng học, thời gian học và thông tin từ vựng đã học.

### 3.2.6. Chức năng “Luyện phát âm”

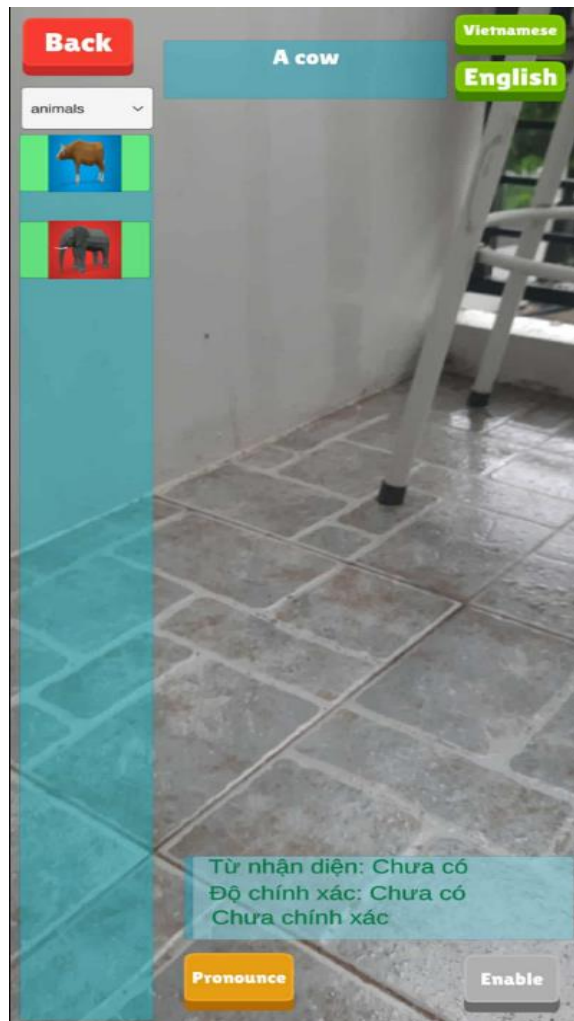
Màn hình chức năng xem luyện phát âm gồm danh sách tùy chọn danh mục mô hình, nút hướng dẫn sử dụng chức năng (nút “?”), nút quay lại (nút “Back”), nút gửi phát âm đã ghi (nút “Pronounce”), nút tùy chọn bật/tắt phát hiện mặt phẳng (nút “Disable”), nút nghe phát âm tiếng anh (nút “English”), nút nghe phát âm tiếng Việt (nút “Vietnamese”).



Hình 3.19: Giao diện hướng dẫn sử dụng chức năng luyện phát âm

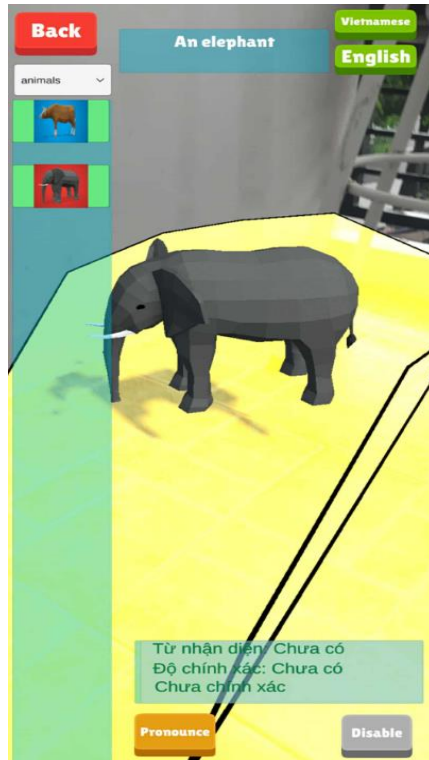
Tương tự với chức năng “Xem thông tin”, chức năng “Luyện phát âm” cũng sẽ có giao diện hướng dẫn sử dụng, có thể chọn danh mục để tải các mô hình muốn học nhưng không cần phải quét hình ảnh chứa thông tin mà chỉ cần quét mặt phẳng để có thể đặt thông tin lên mặt phẳng. Ngoài ra, chức năng sẽ có thêm tính năng ghi âm phát âm của

người dùng và chấm điểm, đồng thời lưu lại các kết quả vào trong bảng tiến độ học tập của người dùng.



Hình 3.20: Giao diện chức năng “Luyện phát âm” sau khi tải thành công mô hình

Trong trường hợp người dùng tải thành công danh sách mô hình từ máy chủ, danh mục bên trái sẽ hiện các mô hình đã tải thành công. Giao diện phụ sẽ hiện từ tiếng Anh của mô hình, điểm phát âm và phát ra âm thanh phát âm của mô hình.



Hình 3.21: Hiển thị mô hình trên mặt phẳng đã được quét trên chức năng “Luyện phát âm”



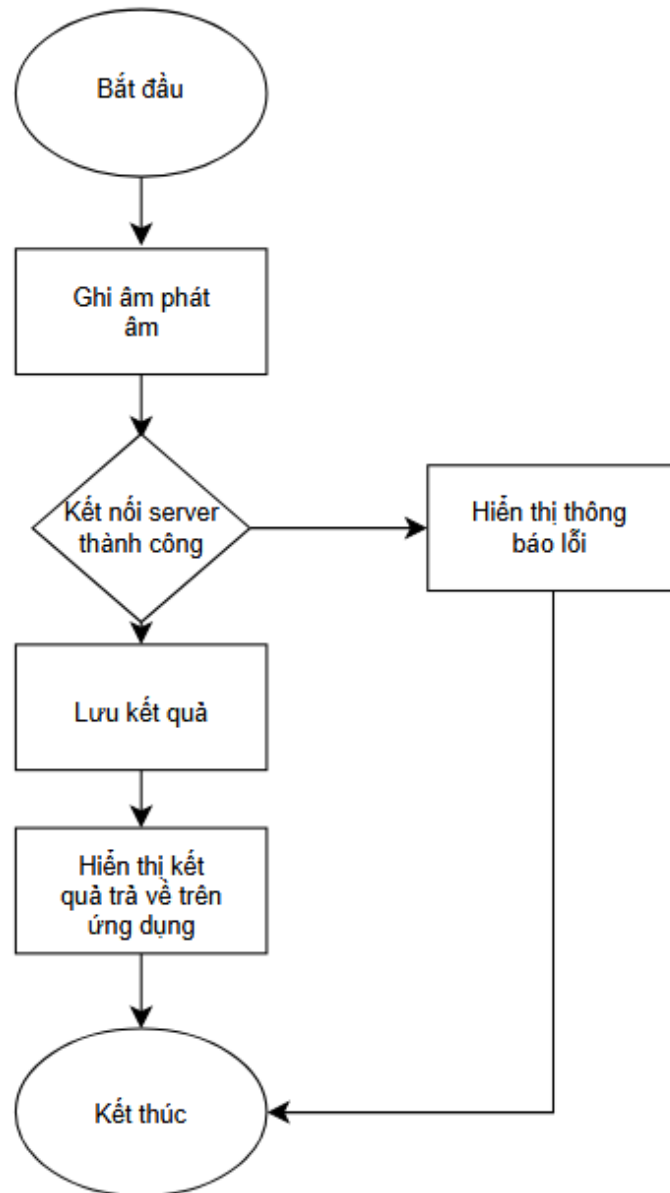
Hình 3.22: Tắt chức năng quét bề mặt phẳng

Khi tắt chức năng quét bề mặt phẳng, chương trình sẽ tạm dừng hàm phát hiện mặt phẳng và tất cả các hình ảnh thể hiện bề mặt phẳng đã được quét. Điều này giúp ứng dụng xử lý nhẹ hơn và tránh việc tràn bộ nhớ. Đồng thời, trải nghiệm người dùng cũng sẽ trở nên tốt và dễ dàng hơn.



Hình 3.23: Minh họa chức năng đánh giá phát âm

Khi người dùng đã chọn mô hình và muốn kiểm tra phát âm của mình, nhấn chọn nút Record (sẽ tự động dừng ghi âm sau 5 giây hoặc dừng sau khi bấm nút lần 2), tệp âm thanh phát âm sẽ được gửi tới server để tách âm thanh thành văn bản, sau đó từ văn bản sử dụng thuật toán Levenshtein để chấm điểm phát âm. Mô tả đơn giản quá trình chấm điểm phát âm như sơ đồ sau.



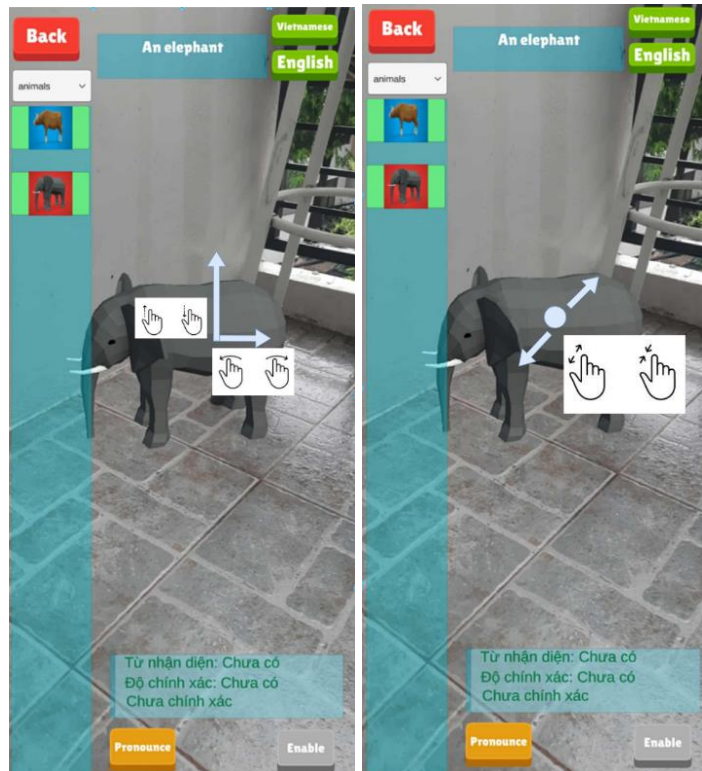
Hình 3.24: Sơ đồ chức năng đánh giá phát âm

### 3.2.7. Chức năng tương tác mô hình

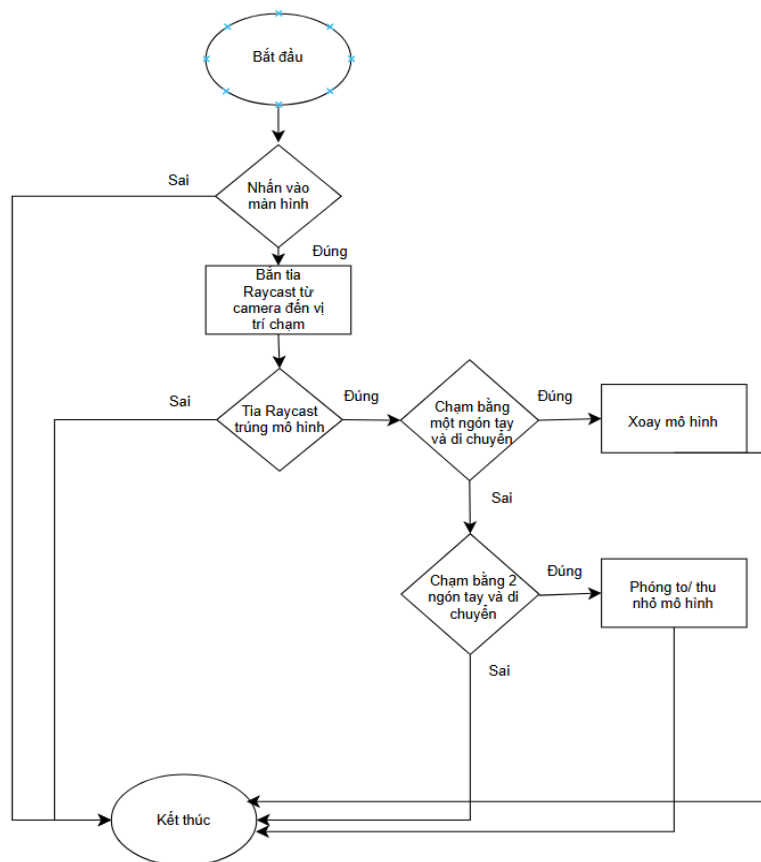
Khi người dùng nhấn vào mô hình 3D, thao tác kéo mô hình sang trái/phải, lên/xuống đến một khoảng nhất định thì mô hình 3D đó sẽ quay sang trái/phải, lên/xuống.

Khi người dùng giữ vào mô hình 3D với 2 ngón tay, thao tác kéo to/thu nhỏ đến một khoảng nhất định thì mô hình 3D đó sẽ được phóng to/thu nhỏ.

Minh họa và sơ đồ hoạt động của những hoạt động tương tác với mô hình được mô tả như sau:







Hình 3.25: Thao tác xoay mô hình trái/phải, lên/xuống và thao tác phóng to/thu nhỏ mô hình



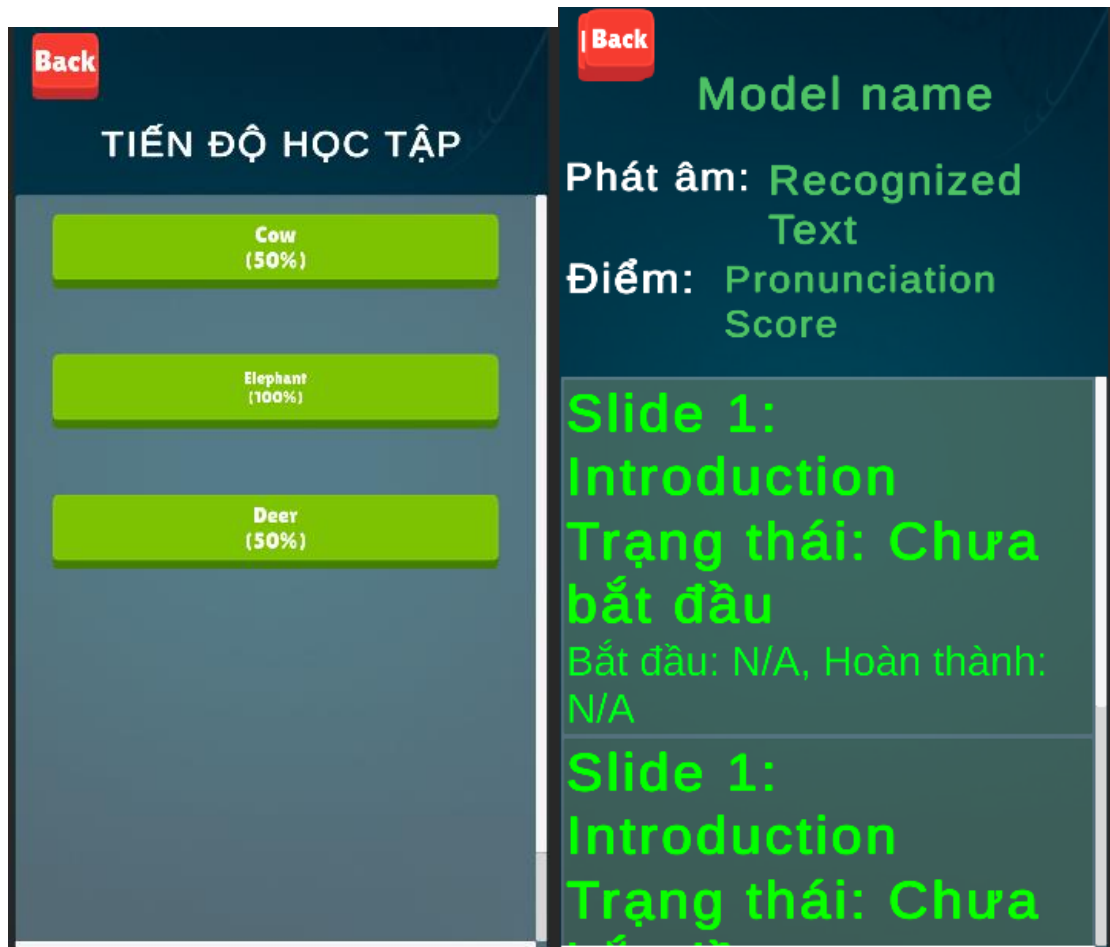
Hình 3.26: Sơ đồ hoạt động chức năng xoay phóng mô hình

## Thao tác của các chức năng tích hợp AR

Bảng 3.4: Bảng danh sách các thao tác để tương tác với mô hình 3D

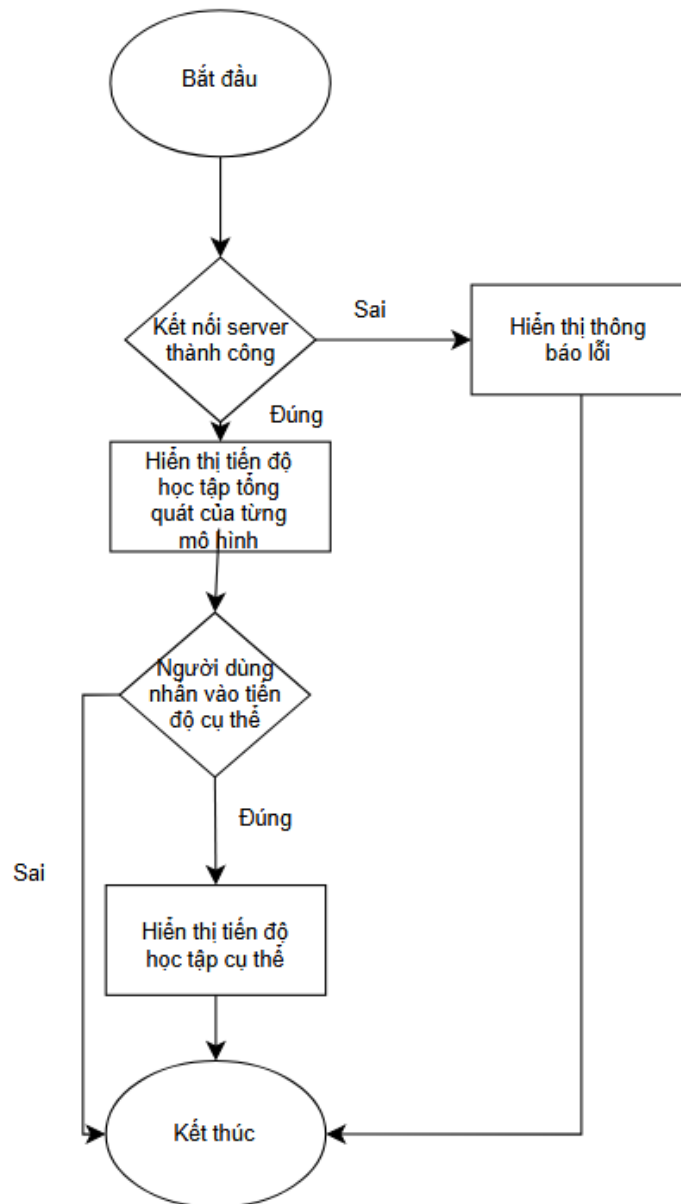
Thao tác	Tên thao tác	Chức năng
	Nhấn	<ul style="list-style-type: none"> <li>• Chọn và đặt mô hình phẳng (đối với chức năng hiển thị mô hình trên mặt phẳng)</li> <li>• Tương tác với thông tin mô hình, nghe phát âm mô hình (đối với chức năng hiển thị mô hình trên chức năng quét hình ảnh)</li> </ul>
	Vuốt sang trái/ phải với một ngón tay	<ul style="list-style-type: none"> <li>• Xoay mô hình được nhấn theo trục Y</li> <li>• Lướt sang trang trình chiếu tiếp theo/ trước đó hiển thị thông tin mô hình</li> </ul>
	Vuốt lên/ xuống với một ngón tay	Xoay mô hình được nhấn theo trục X
	Kéo 2 ngón tay xa ra/ gần lại	Phóng to/ thu nhỏ mô hình

### 3.2.8. Chức năng “Tiền độ học tập”



Hình 3.27: Chức năng xem tiến độ học tập

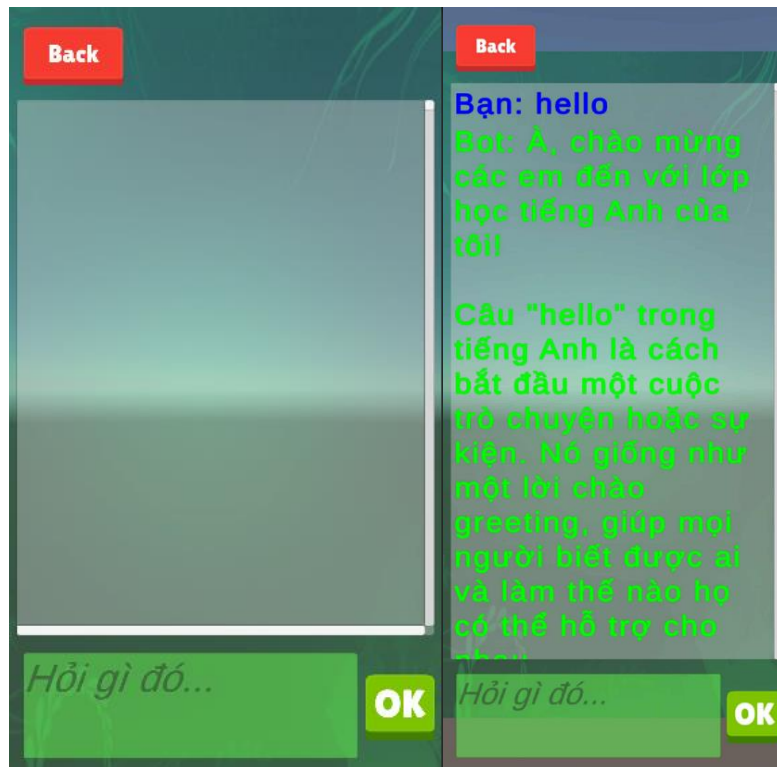
- Tính năng xem tiến độ học tập cho phép xem các tiến độ học tập đối với từng mô hình. Trước tiên các giao diện sẽ hiển thị tiến độ học (dưới dạng nút) tổng quát theo tỉ lệ phần trăm dựa trên các bản trình chiếu của các mô hình.
- Khi người dùng bấm vào từng tiến độ học cụ thể, màn hình sẽ chuyển hướng sang giao diện xem tiến độ học chi tiết gồm tên bản trình chiếu, trạng thái, thời gian bắt đầu, thời gian kết thúc và từ phát âm của người dùng đã được nhận diện thành văn bản cùng với điểm số trung bình của những lần phát âm. Sơ đồ chức năng hiển thị tiến độ học tập được trình bày đơn giản dưới đây.



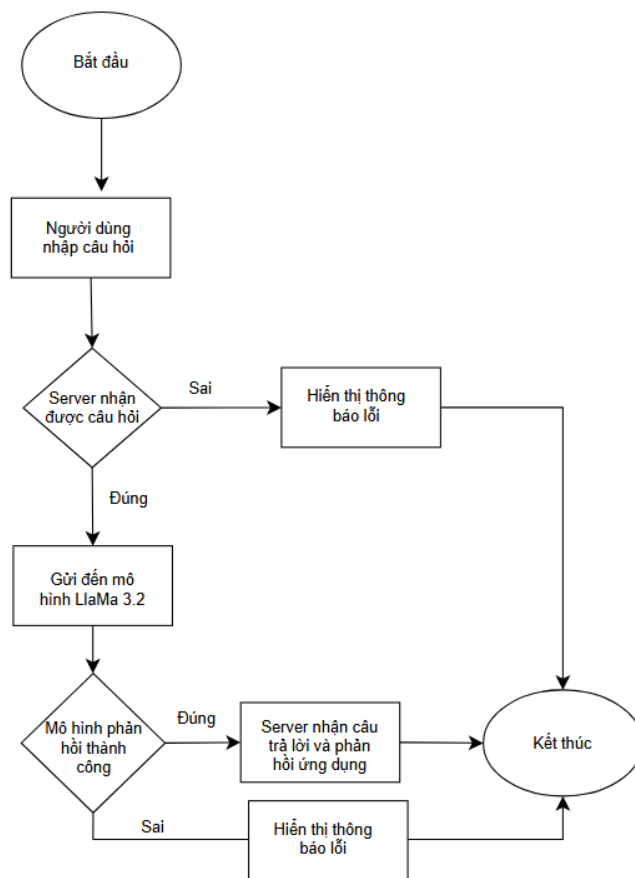
Hình 3.28: Sơ đồ hoạt động chức năng xem tiến độ học tập

### 3.2.9. Chức năng “Trợ lý ảo”

Chức năng trợ lý ảo gồm những thành phần: khung nhập câu hỏi, nút gửi, giao diện hiển thị câu hỏi – câu trả lời và nút quay về màn hình chính. Chức năng này đóng vai trò gọi đến API chatbot của server, server sẽ nhận câu hỏi từ phía ứng dụng và gửi thông tin đến mô hình ngôn ngữ lớn LLaMA 3.2 thông qua Spring AI và Ollama. Khi nhận được câu trả lời, server sẽ truyền lại thông tin nhận được từ mô hình sang ứng dụng. Giao diện và sơ đồ chức năng hoạt động của chức năng “Trợ lý ảo” được trình bày một cách đơn giản như dưới đây:

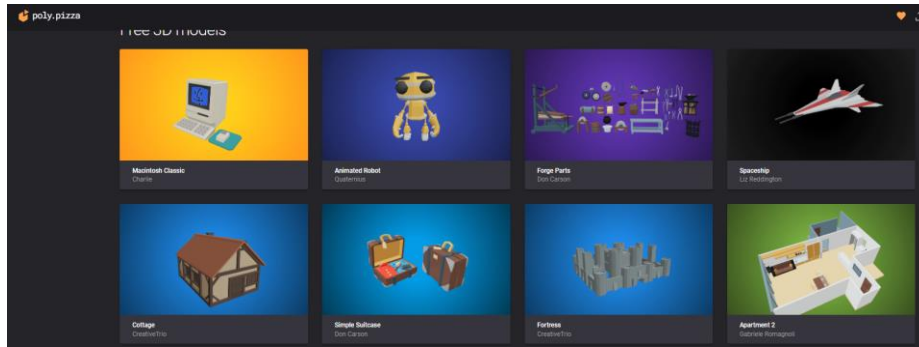


Hình 3.29: Chức năng “Trợ lý ảo”



Hình 3.30: Sơ đồ hoạt động chức năng “Trợ lý ảo”

### 3.3. Mô hình 3D



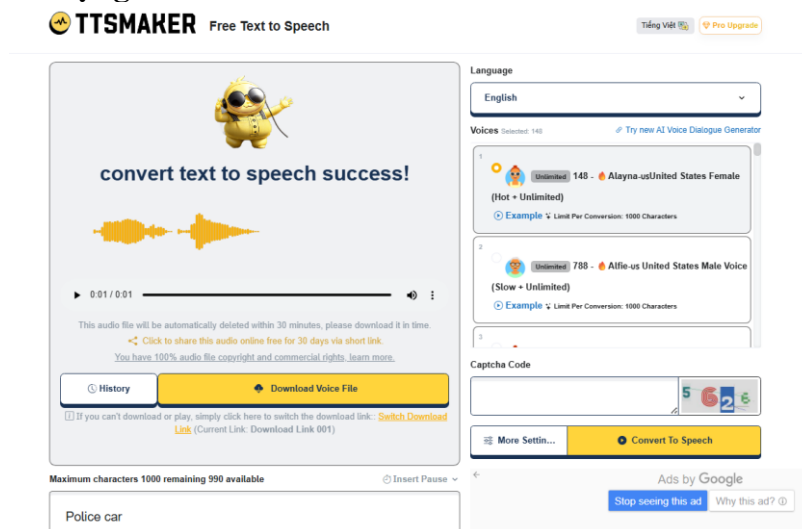
Hình 3.31: Trang web “Pizza poly”

Các mô hình 3D được lấy trên trang web “Piza Poly”.

Poly Pizza (đúng hơn là “Pizza Poly” hoặc “poly.pizza”) là một kho mô hình 3D miễn phí cực kỳ hữu ích cho người làm game, AR/VR, hoặc ứng dụng 3D học tập. Với những lợi ích như sau:

- Miễn phí hoàn toàn: Mô hình có thể được tải xuống và sử dụng miễn phí
- Low-poly: Dạng mô hình ít đa giác
- Dễ dùng: Mô hình tải ở định dạng phổ biến (.glb, .gltf, .fbx, .obj)
- Tải không cần tạo tài khoản

### 3.4. Phát âm từ vựng



Hình 3.32: Trang web “TTS Maker”

Các âm thanh phát âm từ vựng được tạo từ trang web “TTS Maker”, hỗ trợ chuyển từ văn bản sang giọng nói. Dịch vụ của trang web bao gồm nhiều giọng nói, phát âm chuẩn và có bao gồm tiếng Việt

### 3.5. Đánh giá kết quả

Ứng dụng di động hỗ trợ học tiếng anh kết công nghệ AR về cơ bản đã đáp ứng được các tính năng xoay phóng mô hình, tương tác với mô hình giới thiệu các bài học trong không gian thực hỗ trợ gia tăng tính thú vị và trực quan trong học tập đồng thời hỗ trợ giao tiếp với trợ lý ảo, quản lý tiến độ học tập,. Giao diện tối giản, thân thiện, phù hợp với trẻ nhỏ. Tuy nhiên vẫn còn những lỗi về chức năng chương trình và sai sót trong giao diện.

## KẾT LUẬN VÀ HƯỚNG PHÁT TRIỂN

### 1. Kết quả đạt được

Dự án đã hoàn thiện một hệ sinh thái dạy học tiếng Anh ứng dụng công nghệ thực tế tăng cường (AR), nhằm mang đến trải nghiệm học tập sinh động, trực quan và tương tác cao, đặc biệt phù hợp với trẻ em. Thông qua việc kết hợp mô hình 3D minh họa, hệ thống không chỉ giúp cải thiện trí nhớ của trẻ mà còn nâng cao kỹ năng nghe, nói, phát âm. Trong chương trình gồm các chức năng chính:

- Đăng ký, đăng nhập
- Quét hình ảnh có chứa nội dung mô hình và hiển thị mô hình 3D tương ứng
- Hiển thị mô hình trên các bề mặt phẳng được phát hiện
- Trình chiếu thông tin mô hình
- Phát âm từ vựng mô hình chuẩn theo giọng bản ngữ
- Tương tác, xoay phóng mô hình trong không gian thực tế
- Luyện phát âm từ vựng
- Ghi nhận và theo dõi tiến độ học tập của người dùng
- Trò chuyện với trợ lý ảo

### 2. Hạn chế

- Bộ dữ liệu hình ảnh và mô hình hiện tại còn hạn chế, chưa bao phủ toàn bộ các chủ đề từ vựng thông dụng.
- Tính năng nhận diện hình ảnh đôi khi chưa ổn định trong điều kiện ánh sáng yếu hoặc hình ảnh bị mờ.
- Ứng dụng yêu cầu thiết bị có cấu hình tương đối cao để hoạt động mượt mà khi xử lý mô hình AR.
- Chưa có hệ thống bài tập tương tác hoặc game hóa nhằm tăng tính hấp dẫn và duy trì hứng thú học tập lâu dài.

### 3. Hướng phát triển

- Mở rộng cơ sở dữ liệu mô hình 3D và từ vựng với đa dạng chủ đề hơn (động vật, nghề nghiệp, đồ vật, v.v.).
- Tối ưu thuật toán nhận diện hình ảnh để tăng độ chính xác và tốc độ xử lý.
- Tích hợp tính năng học qua trò chơi (gamification) để nâng cao trải nghiệm người dùng.
- Xây dựng phiên bản dành cho giáo viên với chức năng quản lý lớp học và đánh giá học viên.
- Phát triển mô-đun học theo chủ đề và lộ trình cá nhân hóa cho từng trình độ người học.
- Hỗ trợ đa nền tảng (iOS, Android, Web AR) để mở rộng đối tượng sử dụng.

## TÀI LIỆU THAM KHẢO

- [1]. Joe Hocking, “*Unity in Action, Third Edition: Multiplatform game development in C# 3rd ed. Edition*”, Manning Publication. (8/2/2022).
- [2]. Google, “*ARCore*”, Google Developer (31/10/2022).  
<https://developers.google.com/ar/discover>
- [3]. Sebastián Peyrott., “*JWT Handbook*”, Auth0 Inc (2016).
- [4]. Craig Walls, “*Spring Boot in Action First Edition*, Manning Publication (3/1/2016).
- [5]. Unity, “*AR Foundation*”, Unity Technologies. (2018).  
<https://docs.unity3d.com/Packages/com.unity.xr.arfoundation@6.1/manual/index.html>
- [6]. Daniel Nichter, “*Efficient MySQL Performance: Best Practices and Techniques 1st Edition*”. O'Reilly Media Publication (4/1/2022).
- [7]. Spring, “*Spring Boot*”, Broadcomm Inc (n.d).  
<https://docs.spring.io/spring-boot/index.html>
- [8]. AI at Meta, “*Llama 3.2*”, Meta (n.d).  
[https://www.llama.com/docs/model-cards-and-prompt-formats/llama3\\_2/](https://www.llama.com/docs/model-cards-and-prompt-formats/llama3_2/)
- [9]. SYSTRAN, “*Faster Whisper*”, Systran Soft (n.d).  
<https://github.com/SYSTRAN/faster-whisper>

