# ĐẠI HỌC QUỐC GIA TP. HỒ CHÍ MINH

**TRƯỜNG ĐẠI HỌC CÔNG NGHỆ THÔNG TIN KHOA CÔNG NGHỆ PHẦN MỀM**

# LÊ TRỌNG TÍN – 15520893 BÙI PHẠM MINH THI - 15520824

**KHÓA LUẬN TỐT NGHIỆP**

XÂY DỰNG ỨNG DỤNG NHẬN DẠNG CHỮ VIẾT VÀ BIỂU DIỄN MÔ HÌNH AR

Application of recognizing text and performing AR model

# KỸ SƯ NGÀNH KỸ THUẬT PHẦN MỀM

**ĐẠI HỌC QUỐC GIA TP. HỒ CHÍ MINH**

**TRƯỜNG ĐẠI HỌC CÔNG NGHỆ THÔNG TIN KHOA CÔNG NGHỆ PHẦN MỀM**

# LÊ TRỌNG TÍN – 15520893 BÙI PHẠM MINH THI - 15520824

**KHÓA LUẬN TỐT NGHIỆP**

XÂY DỰNG ỨNG DỤNG NHẬN DẠNG CHỮ VIẾT VÀ BIỂU DIỄN MÔ HÌNH AR

Application of recognizing text and performing AR model

# KỸ SƯ NGÀNH KỸ THUẬT PHẦN MỀM

**GIẢNG VIÊN HƯỚNG DẪN TS. NGUYỄN THỊ THANH TRÚC**

## DANH SÁCH HỘI ĐỒNG BẢO VỆ KHÓA LUẬN

Hội đồng chấm khóa luận tốt nghiệp, thành lập theo Quyết định số ……………………

ngày của Hiệu trưởng Trường Đại học Công nghệ Thông tin.

1. Chủ tịch.
2. Thư ký.
3. Ủy viên.
4. Ủy viên.

ĐHQG TP. HỒ CHÍ MINH **CỘNG HOÀ XÃ HỘI CHỦ NGHĨA VIỆT NAM**

**TRƯỜNG ĐẠI HỌC Độc Lập – Tự Do – Hạnh Phúc CÔNG NGHỆ THÔNG TIN** TP. HCM, ngày 01 tháng 12 năm 2019

## NHẬN XÉT KHOÁ LUẬN TỐT NGHIỆP (CỦA CÁN BỘ HƯỚNG DẪN)

**Tên khoá luận:**

**XÂY DỰNG ỨNG DỤNG NHẬN DẠNG CHỮ VIẾT VÀ BIỂU DIỄN MÔ HÌNH AR**

|  |  |  |
| --- | --- | --- |
| **Nhóm SV thực hiện:**  Lê Trọng Tín | 15520893 | **Cán bộ hướng dẫn:**  TS. Nguyễn Thị Thanh Trúc |
| Bùi Phạm Minh Thi | 15520824 |  |
| **Đánh giá khoá luận:**  1. Về cuốn báo cáo: Số trang: |  | Số chương: |
| Số bảng số liệu:  Số tài liệu tham khảo: |  | Số hình vẽ:  Sản phẩm: |

Một số nhận xét về hình thức cuốn báo cáo:

································ ································ ································

································ ································ ································

································ ································ ································

1. Về nội dung nghiên cứu:

|  |  |  |  |
| --- | --- | --- | --- |
| ································ | ································ | ································ | ·· |
| ································ | ································ | ································ | ·· |

|  |  |  |  |
| --- | --- | --- | --- |
| ································ | ································ | ································ | ·· |
| ································ | ································ | ································ | ·· |

1. Về chương trình ứng dụng:

································ ································ ································ ··

································ ································ ································ ··

································ ································ ································ ··

································ ································ ································ ··

1. Về thái độ làm việc của sinh viên:

································ ································ ································ ··

································ ································ ································ ··

································ ································ ································ ··

**Đánh giá chung:**

································ ································ ································ ··

································ ································ ································ ··

································ ································ ································ ··

································ ································ ································ ··

**Điểm từng sinh viên:**

Lê Trọng Tín: **/10**

Bùi Phạm Minh Thi: **/10**

## Người nhận xét

(Ký tên và ghi rõ họ tên)

## TS. Nguyễn Thị Thanh Trúc

ĐHQG TP. HỒ CHÍ MINH **CỘNG HOÀ XÃ HỘI CHỦ NGHĨA VIỆT NAM**

**TRƯỜNG ĐẠI HỌC Độc Lập – Tự Do – Hạnh Phúc CÔNG NGHỆ THÔNG TIN** TP. HCM, ngày 01 tháng 12 năm 2019

## NHẬN XÉT KHOÁ LUẬN TỐT NGHIỆP (CỦA CÁN BỘ PHẢN BIỆN)

**Tên khoá luận:**

**XÂY DỰNG ỨNG DỤNG NHẬN DẠNG CHỮ VIẾT VÀ BIỂU DIỄN MÔ HÌNH AR**

**Nhóm SV thực hiện: Cán bộ phản biện:**

Lê Trọng Tín 15520893

Bùi Phạm Minh Thi 15520824

## Đánh giá khoá luận:

1. Về cuốn báo cáo:

Số trang: Số chương:

Số bảng số liệu: Số hình vẽ:

Số tài liệu tham khảo: Sản phẩm: Một số nhận xét về hình thức cuốn báo cáo:

|  |  |  |
| --- | --- | --- |
| ································ ································ | ································ |  |
| ································ ································ | ································ |
| ································ ································ | ································ |
| 2. Về nội dung nghiên cứu: |  |
| ································ ································ | ································ | ·· |
| ································ ································ | ································ | ·· |

|  |  |  |  |
| --- | --- | --- | --- |
| ································ | ································ | ································ | ·· |
| ································ | ································ | ································ | ·· |

* 1. Về chương trình ứng dụng:

································ ································ ································ ··

································ ································ ································ ··

································ ································ ································ ··

································ ································ ································ ··

* 1. Về thái độ làm việc của sinh viên:

································ ································ ································ ··

································ ································ ································ ··

································ ································ ································ ··

**Đánh giá chung:**

································ ································ ································ ··

································ ································ ································ ··

································ ································ ································ ··

································ ································ ································ ··

**Điểm từng sinh viên:**

Lê Trọng Tín: **/10**

Bùi Phạm Minh Thi: **/10**

## Người nhận xét

(Ký tên và ghi rõ họ tên)

## LỜI CẢM ƠN

Sau 4 năm học tập và rèn luyện tại Trường Đại Học Công Nghệ Thông Tin, được sự chỉ bảo và giảng dạy nhiệt tình của quý thầy cô, đặc biệt quý thầy cô trong khoa Công Nghệ Phần Mềm, chúng em đã nhận được vô vàng kiến thức bổ ích về lý thuyết và thực hành trong thời gian qua. Trong khoảng thời gian thực hiện luận văn tốt nghiệp, chúng em càng học hỏi thêm được nhiều kiến thức, kinh nghiệm để áp dụng và đã hoàn thành được luận văn tốt nghiệp của chúng em.

Xin chân thành cảm ơn quý thầy cô Khoa Công Nghệ Phần mềm, đặc biệt là cô Nguyễn Thị Thanh Trúc đã tận tình hướng dẫn chúng em hoàn thành luận văn tốt nghiệp này.

Bên cạnh đó, một lần nữa xin cảm ơn nhóm các bạn cùng lớp làm khóa luận đã động viên, thảo luận và góp ý cho nhóm đồng thời đã khơi thêm nguồn động lực cho nhóm để tiếp thêm động lực cho nhóm trong giai đoạn đầy khó khăn.

Ngoài ra, chúng em cũng cảm ơn các anh, chị, bạn bè và gia đình đã giúp đỡ và tạo điều kiện cho chúng em hoàn tất luận văn này.

Mặc dù đã rất cố gắng, tuy nhiên vì khả năng giới hạn nên vẫn không tránh khỏi những sai sót, chúng em rất mong nhận được sự thông cảm và góp ý chân tình từ quý thầy cô và các bạn. Một lần nữa xin được chân thành cảm ơn mọi người.

Thành phố Hồ Chí Minh, tháng 12 năm 2019

Lê Trọng Tín Bùi Phạm Minh Thi

ĐHQG TP. HỒ CHÍ MINH **CỘNG HOÀ XÃ HỘI CHỦ NGHĨA VIỆT NAM**

## TRƯỜNG ĐẠI HỌC Độc Lập – Tự Do – Hạnh Phúc CÔNG NGHỆ THÔNG TIN

**ĐỀ CƯƠNG CHI TIẾT**

|  |
| --- |
| **TÊN ĐỀ TÀI: ỨNG DỤNG NHẬN DẠNG CHỮ VIẾT VÀ BIỂU DIỄN MÔ**  **HÌNH AR** |
| **Cán bộ hướng dẫn:** ThS. Nguyễn Thị Thanh Trúc |
| **Thời gian thực hiện:** Từ ngày 01/09/2019 đến ngày 31/12/2019 |
| **Sinh viên thực hiện:**  Bùi Phạm Minh Thi – 15520824 Lê Trọng Tín – 15520893 |
| **Nội dung đề tài:**  ***Mục tiêu***:   * Xây dựng ứng dụng nhận dạng chữ viết với đa ngôn ngữ. * Xây dựng ứng dụng biểu diễn mô hình thực tế ảo tăng cường tương ứng với chữ viết nhận nhận diện được bằng công nghệ AR. * Ứng dụng có khả năng mở rộng và tương tác cao. * Ứng dụng có giao diện trực quan, thân thiện phù hợp đối tượng người dùng.   ***Phạm vi***:  Phạm vi địa lý:Việt Nam. Phạm vi chức năng:   * Nhận diện chữ chính xác. * Xử lý với thời gian thực. * Hỗ trợ đa ngôn ngữ. |

|  |
| --- |
| * Biểu diễn mô hình thực tế ảo tăng cường.   ***Đối tượng sử dụng***:  Trẻ em, học sinh, sinh viên, những người đam mê tìm hiểu các công nghệ mới. Đối tượng người dùng tập trung vào giới trẻ từ độ tuổi 10 đến 30.  ***Phương pháp thực hiện***:   * Ứnng dụng được xây dựng trên nền tảng Android. * Tiền xử lý ảnh bằng thư viện openCV. * Sử dụng công nghệ nhận dạng ký tự quang học Tesseract do Google phát triển để nhận dạng chữ viết. * Nhận dạng chữ viết với thời gian thực. * Ứng dụng hỗ trợ nhận dạng đa ngôn ngữ. * Sử dụng bộ training data dựa trên phương pháp Long Short Term Memory (LSTM). * Phát triển ứng dụng kết hợp biểu diễn mô hình thực tế ảo tăng cường từ kết quả nhận dạng chữ bằng gói ARCore - nền tảng trải nghiệm thực tế ảo tăng cường của Google.   ***Kết quả mong đợi:***   * Xây dựng ứng dụng dễ dàng sử dụng, hấp dẫn với tính năng mới mẽ. * Ứng dụng giúp ích người dùng trong việc học tập và nghiên cứu công nghệ mới. * Ứng dụng đạt đủ tiêu chí và yêu cầu đặt ra. |
| **Kế hoạch thực hiện**  ***Kế hoạch làm việc:***   * Mỗi tuần liên lạc với giảng viên một lần để trình bày kết quả, nghe nhận xét và góp ý từ đó để đưa ra kế hoạch tiếp theo. * Kế hoạch thực hiện: 16 tuần. Bắt đầu vào 5/9/2019 và kết thức 31/12/2019. |

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | **Giai đoạn** | **Thời gian** | **Công việc** |  |
| **Giai đoạn 1: Khảo sát,**  **nghiên cứu** | Tuần 1 (5/9 – 12/9) | Giới thiệu đề tài. Phát biểu bài toán.  Khảo sát thị trường. |
| **Giai đoạn 2: Chuẩn bị** | Tuần 2 (13/9 – 23/9) Tuần 3 (24/9 – 30/9) Tuần 4 (1/10 – 7/10) | Tìm hiểu về các công nghệ áp dụng và các thuật sử dụng.  Phân tích các yêu cầu chức năng của ứng dụng.  Thiết kế sơ đồ tổng quan.  Thiết lập môi trưởng triển khai. |
| **Giai đoạn 3: Triển khai** | Tuần 5 (8/10 – 14/10) Tuần 6 (15/10 – 21/10) Tuân 7 (22/10 – 28/10) | Tiền xử lý ảnh với openCV.  Nhận dạng chữ viết với Tesseract. Phát triển ứng dụng xử lý với thời gian thực. |
| **Giai đoạn 4: Triển khai** | Tuần 8 (29/10 – 4/11) Tuần 9 (5/11 – 11/11) Tuần 10 (12/11 – 18/11) | Xây dựng chức năng chuyển đổi ngôn ngữ nhận dạng.  Cài đặt và sử dụng bộ training dữ  liệu hiệu quả. |
| **Giai đoạn 5: Triển khai** | Tuần 11 (19/11 – 25/11)  Tuần 12 (26/11 – 2/12) Tuần 13 (3/12 – 9/12) | Sử dụng ARCore để biểu diễn mô hình thực tể ảo tăng cường. |
| **Giai đoạn 6:**  **Kiểm thử** | Tuần 14 (10/12 – 16/12)  Tuần 15 (17/12 – 23/12) | Kiểm thử.  Thu thập kết quả. |
| **Giai đoạn 7:**  **Tổng kết** | Tuần 16 (24/12 – 31/12) | Hoàn thành báo cáo. |
|  | | | | |

**Xác nhận của CBHD**

(Ký tên và ghi rõ họ tên)

TS. Nguyễn Thị Thanh Trúc

|  |  |
| --- | --- |
| Bùi Phạm Minh Thi | Lê Trọng Tín |

## TP. HCM, ngày tháng năm 2019 Sinh viên

(Ký tên và ghi rõ họ tên)

## MỤC LỤC

[CHƯƠNG 1. TỔNG QUAN 1](#_bookmark0)

* 1. [Đặt vấn đề 1](#_bookmark1)
  2. [Mục tiêu và phạm vi 1](#_bookmark2)
     1. [Mục tiêu 1](#_bookmark3)
     2. [Phạm vi 2](#_bookmark4)
  3. [Nội dung thực hiện 2](#_bookmark5)
  4. [Bố cục báo cáo 2](#_bookmark6)

[CHƯƠNG 2. PHÂN TÍCH XÁC ĐINH BÀI TOÁN 4](#_bookmark7)

* 1. [Khảo sát thực trạng 4](#_bookmark8)
  2. [Các hệ thống liên quan 7](#_bookmark13)
     1. [AR animal 8](#_bookmark14)
     2. [AR Animals 9](#_bookmark16)
     3. [Planet AR - Animals 10](#_bookmark18)
     4. [Tổng hợp 11](#_bookmark20)
  3. [Yêu cầu chức năng 11](#_bookmark21)
  4. [Yêu cầu phi chức năng 12](#_bookmark22)

[CHƯƠNG 3. TÌM HIỂU VÀ LỰA CHỌN CÔNG NGHỆ 13](#_bookmark23)

* 1. [Tổng quan về các công nghệ trong ứng dụng 13](#_bookmark24)
     1. [Tesseract OCR 13](#_bookmark25)
     2. [AR – AR Core 28](#_bookmark44)
     3. [Các công nghệ khác 36](#_bookmark51)
  2. [Kết hợp các công nghệ để xây dựng ứng dụng 40](#_bookmark54)

[CHƯƠNG 4. PHÂN TÍCH VÀ THIẾT KẾ HỆ THỐNG 41](#_bookmark56)

* 1. [Phân tích yêu cầu 41](#_bookmark57)
     1. [Sơ đồ use case 41](#_bookmark58)
     2. [Danh sách actor 42](#_bookmark60)
     3. [Danh sách use case 42](#_bookmark62)
     4. [Mô tả chi tiết các use case 43](#_bookmark64)
  2. [Thiết kế dữ liệu 58](#_bookmark83)
     1. [Sơ đồ các thành phần chính 58](#_bookmark84)
     2. [Mô tả các thành phần có trong ứng dụng 58](#_bookmark86)
     3. [Danh sách các ngôn ngữ nhận dạng 60](#_bookmark88)
     4. [Danh sách các động vật 63](#_bookmark90)
  3. [Thiết kế kiến trúc 67](#_bookmark92)
     1. [Sơ đồ kiến trúc tổng thể 67](#_bookmark93)
     2. [Sơ đồ quản lý trạng thái 68](#_bookmark95)
     3. [Sơ đồ kiến trúc chi tiết 69](#_bookmark97)
  4. [Thiết kế giao diện 70](#_bookmark99)
     1. [Sơ đồ thiết kế giao diện 70](#_bookmark100)
     2. [Danh sách giao diện 70](#_bookmark102)
     3. [Mô trả chi tiết giao diện 71](#_bookmark104)

[CHƯƠNG 5. KẾT LUẬN VÀ HƯỚNG PHÁT TRIỂN 82](#_bookmark128)

* 1. [Kết quả đạt được 82](#_bookmark129)
  2. [Nhận xét 82](#_bookmark130)
     1. [Ưu điểm 82](#_bookmark131)
     2. [Khuyết điểm 82](#_bookmark132)
  3. [Hướng phát triển 83](#_bookmark133)

[TÀI LIỆU THAM KHẢO 84](#_bookmark134)

**DANH MỤC HÌNH**

[Hình 2.1 - Độ lớn thị trường của AR và VR trong vài năm tới 5](#_bookmark9)

[Hình 2.2 - Số lượng thiết bị bán ra trong vài năm tới 6](#_bookmark10)

[Hình 2.3 - Độ lớn thị trường của các thiết bị AR và VR vào năm 2020 và 2025 6](#_bookmark11)

[Hình 2.4 - Số thiết bị bán ra theo từng hãng 7](#_bookmark12)

[Hình 2.5 - Ứng dụng AR animal 8](#_bookmark15)

[Hình 2.6 - Ứng dụng AR Animals 9](#_bookmark17)

[Hình 2.7 - Ứng dụng Plant AR - Animals 10](#_bookmark19)

[Hình 3.1 - Tesseract OCR 13](#_bookmark26)

[Hình 3.2 - Flow cơ bản của Tesseract 14](#_bookmark27)

[Hình 3.3 - Cơ chế hoạt động của Tesseract 15](#_bookmark28)

Hình 3.4 - Các khối blob sau khi tách 16

[Hình 3.5 - Cấu trúc của Tesseract 17](#_bookmark29)

[Hình 3.6 - Cơ chế nhận dạng kí tự 17](#_bookmark30)

Hình 3.7 - Các khả năng của chữ “O” 18

[Hình 3.8 - Các phân đoạn đoạn chữ “O” 18](#_bookmark31)

Hình 3.9 - Cơ chế chuyển đổi nhận dạng 18

Hình 3.10 - Cơ chế chuyển đổi nhận dạng 19

[Hình 3.11 - Cơ chế Deep Belief nets 19](#_bookmark32)

[Hình 3.12 - Cơ chế độ dốc 20](#_bookmark33)

[Hình 3.13 - Cơ chế chuẩn hóa phân loại 21](#_bookmark34)

[Hình 3.14 - Cơ chế chuẩn hóa phân loại 21](#_bookmark35)

[Hình 3.15 - Mạng nơ-ron hồi quy 23](#_bookmark37)

[Hình 3.16 - Mạng nơ-ron hồi quy 23](#_bookmark38)

[Hình 3.17 - Mạng nơ-ron hồi quy 24](#_bookmark39)

[Hình 3.18 - Mạng nơ-ron hồi quy 24](#_bookmark40)

[Hình 3.19 - LSTM cơ bản 26](#_bookmark41)

[Hình 3.20 - LSTM với softmax 26](#_bookmark42)

[Hình 3.21 - Tesseract sử dụng LSTM 27](#_bookmark43)

[Hình 3.22 - Công nghê AR 28](#_bookmark45)

[Hình 3.23 - AR Core 31](#_bookmark46)

[Hình 3.24 - Cơ chế hoạt động ARCore 32](#_bookmark47)

[Hình 3.25 - Cơ chế theo dõi chuyển động 34](#_bookmark48)

[Hình 3.26 - Cơ chế nhận biết môi trường 35](#_bookmark49)

[Hình 3.27 - Cơ chế ước lượng ánh sáng 35](#_bookmark50)

[Hình 3.28 - Hệ điều hành Android 36](#_bookmark52)

[Hình 3.29 - Thư viện OpenCV 39](#_bookmark53)

[Hình 3.30 - Các công nghệ trong ứng dụng 40](#_bookmark55)

[Hình 4.1 - Mô hình use case ứng dụng 41](#_bookmark59)

[Hình 4.2 - Sơ đồ các thành phần chính 58](#_bookmark85)

[Hình 4.3 - Mô hình kiến trúc tổng thể 67](#_bookmark94)

[Hình 4.4 - Sơ đồ hoạt động nhận dạng 68](#_bookmark96)

[Hình 4.5 - Sơ đồ kiến trúc chi tiết 69](#_bookmark98)

[Hình 4.6 - Sơ đồ thiết kế giao diện 70](#_bookmark101)

[Hình 4.7 - Giao diện “Khởi động” 71](#_bookmark105)

[Hình 4.8 - Giao diện “Giới thiệu tính năng” 72](#_bookmark107)

[Hình 4.9 - Giao diện “Giới thiệu tính năng” 72](#_bookmark108)

[Hình 4.10 - Giao diện “Giới thiệu tính năng” 73](#_bookmark109)

[Hình 4.11 - Giao diện “Giới thiệu tính năng” 73](#_bookmark110)

[Hình 4.12 - Giao diện “Home” 74](#_bookmark112)

[Hình 4.13 - Giao diện nhận dạng chữ 74](#_bookmark114)

[Hình 4.14 - Giao diện “Danh sách động vật” 75](#_bookmark116)

[Hình 4.15 - Giao diện “Câu hỏi” 76](#_bookmark118)

[Hình 4.16 - Giao diện “Dò mặt phẳng” 77](#_bookmark120)

[Hình 4.17 - Giao diện “Tương tác mô hình” 78](#_bookmark122)

[Hình 4.18 - Giao diện “Hướng dẫn đặt mô hình” 79](#_bookmark124)

[Hình 4.19 - Giao diện “Cài đặt” 80](#_bookmark126)

## DANH MỤC BẢNG

[Bảng 3.1 - Bảng so sánh các công cụ nhận dạng chữ 22](#_bookmark36)

[Bảng 4.1 - Danh sách actor 42](#_bookmark61)

[Bảng 4.2 - Dang sách use case 43](#_bookmark63)

[Bảng 4.3- Dòng sự kiện chính use case “Xem giới thiệu tính năng” 44](#_bookmark65)

[Bảng 4.4 - Dòng sự kiện chính use case “Nhận dạng ký tự” 45](#_bookmark66)

[Bảng 4.5 - Dòng sự kiện phụ use case “Nhận dạng ký tự” 46](#_bookmark67)

[Bảng 4.6 - Dòng sự kiện chính use case “Xem danh sách các đông vật” 46](#_bookmark68)

[Bảng 4.7- Dòng sự kiện chính use case "Trả lời câu hỏi" 47](#_bookmark69)

[Bảng 4.8 - Dòng sự kiện phụ use case “Trả lời câu hỏi” 48](#_bookmark70)

[Bảng 4.9 - Dòng sự kiện chính use case “Dò tìm mặt phẳng” 49](#_bookmark71)

[Bảng 4.10 - Dòng sự kiện phụ use case “Dò tìm mặt phẳng” 49](#_bookmark72)

[Bảng 4.11 - Dòng sự kiện chính use case “Đặt mô hình lên mặt phẳng” 50](#_bookmark73)

[Bảng 4.12 - Dòng sự kiện chính use case “Tương tác mô hình” 51](#_bookmark74)

[Bảng 4.13 - Dòng sự kiện chính use case “Xem hướng dẫn đặt mô hình” 52](#_bookmark75)

[Bảng 4.14 - Dòng sự kiện chính use case “Thiết lập hiện thông tin nhận dạng chữ” 53](#_bookmark76)

[Bảng 4.15 - Dòng sự kiện chính use case “Thiết lập thay đổi ngôn ngữ nhận dạng” 54](#_bookmark77)

[Bảng 4.16 - Dòng sự kiện phụ use case “Thiết lập thay đổi ngôn ngữ nhận dạng” 54](#_bookmark78)

[Bảng 4.17 - Dòng sự kiện chính use case “Thiết lập chế độ auto focus” 55](#_bookmark79)

[Bảng 4.18 - Dòng sự kiện chính use case “Thiết lập chế độ focus tiêu chuẩn” 56](#_bookmark80)

[Bảng 4.19 - Dòng sự kiện chính use case “Thiết lập tiếng thông báo” 57](#_bookmark81)

[Bảng 4.20 - Dòng sự kiện chính use case “Thiết lập âm thanh” 57](#_bookmark82)

[Bảng 4.21 - Các thành phần chính trong cơ sở ứng dụng 58](#_bookmark87)

[Bảng 4.22 - Danh sách các ngôn ngữ nhận dạng 63](#_bookmark89)

[Bảng 4.23 - Danh sách các động vật 67](#_bookmark91)

[Bảng 4.25 - Danh sách các giao diện 71](#_bookmark103)

[Bảng 4.26 - Thành phần giao diện “Khởi động” 72](#_bookmark106)

[Bảng 4.27 - Thành phần giao diện “Giới thiệu tính năng” 73](#_bookmark111)

[Bảng 4.28 – Thành phần giao diện “Home” 74](#_bookmark113)

[Bảng 4.29 – Thành phần giao diện “Nhận dạng chữ” 75](#_bookmark115)

[Bảng 4.30 – Thành phần giao diện “Danh sách động vật” 76](#_bookmark117)

[Bảng 4.31 – Thành phần giao diện “Câu hỏi” 76](#_bookmark119)

[Bảng 4.32 – Thành phần giao diện “Dò mặt phẳng” 77](#_bookmark121)

[Bảng 4.33 - Thành phần giao diện “Tương tác mô hình” 78](#_bookmark123)

[Bảng 4.34 - Thành phần giao diện “Hướng dẫn đặt mô hình” 79](#_bookmark125)

[Bảng 4.35 - Thành phần giao diện “Hướng dẫn đặt mô hình” 81](#_bookmark127)

# CHƯƠNG 1. TỔNG QUAN

## Đặt vấn đề

Ngày xưa, việc học tên các loài động vật và tìm hiểu về đời sống của chúng chỉ có thể thực hiện qua sách vở đơn thuần. Hiện đại hơn một chút, những năm gần đây đã có Flashcards với hình ảnh tươi mới và dễ tương tác trên tay hơn nhưng nhìn chung vẫn khá nhàm chán với trẻ em. Tuy nhiên với sự phát triển của công nghệ, mọi thứ đã không còn tẻ nhạt mà trở nên sống động hơn nhiều. Cụ thể, đó là nhờ công nghệ thực tế ảo tăng cường AR được tích hợp trên rất nhiều điện thoại Smartphone này ngay, đem lại cho người dùng chúng ta những trải nghiệm rất vui nhộn, thú vị và không kém phần hữu ích. Vấn đề là các ứng dụng các ứng dụng hiện tại đòi hỏi phải sở hữu bộ card do nhà phát hành cung cấp và phải chi một một khoản tiền kha khá để có được chúng. Ngoài ra các mô hình con vật chỉ xuất hiện trên những tấm card bằng giấy làm cho trải nghiệm ít đi phần thú vị.

Với mục đích tạo trẻ em có cái nhìn thực tế hơn về thế giới động vật và giúp ích cho việc giáo dục, nhóm sẽ tạo ra ứng dụng AR Animal kết hợp việc nhận dạng tên con vật và công nghệ thực tế ảo AR. Bằng cách sử dụng 1 điện thoại hoặc 1 máy tính bảng để quét tên con vật, chọn mặt phẳng để đặt chúng trên màn hình và sau đó con vật sẽ xuất hiện một cách sinh động. Không chỉ được quan sát muông thú một cách gần gũi như ở ngoài thực tế mà còn cung cấp đủ mọi thông tin về loài vật ấy nữa. Bọn trẻ cũng có thể xoay vòng, phóng to hay thu nhỏ kích cỡ của những "người bạn thú ảo" theo ý thích.

## Mục tiêu và phạm vi

## Mục tiêu

* Ứng dụng có thể góp ích trong việc giáo dục, tăng tính sáng tạo và phát triển tư duy trẻ em.
* Người dùng có thêm một ứng dụng để trải nghiệm thực tế ảo tăng cường.
* Hỗ trợ nhận dạng chữ với đa ngôn ngữ.
* Ứng dụng có khả năng mở rộng và tương tác cao, giao diện sinh động, dễ sử dụng phù hợp đối tượng người dùng.

## Phạm vi

* Tiếp cận độ tuổi trẻ em, học sinh, sinh viên.
* Những ai đam mê sáng tạo và công nghệ.
* Các các loài động vật trong và ngoài nước.
* Lựa chọn các công nghệ mới, phổ biến và phù hợp để xây dựng ứng dụng.

## Nội dung thực hiện

* Phân tích đề tài.
* Phân tích bài toán, lên kế hoạch, tìm hiểu các công nghệ sử dụng cho việc xây dựng và phát triển ứng dụng.
* Xác định, đặc tả chức năng của ứng dụng.
* Phát triển sản phẩm qua các giai đoạn.
* Kiểm thử và đánh giá hoàn thiện ứng dụng.
* Triển khai ứng dụng trong thực tế.
* Viết báo cáo luận văn.

## Bố cục báo cáo

Chương 1: TỔNG QUAN: Trình bày về phần đặt vấn đề, lý do chọn đề tài, mục tiêu và phạm vi tiếp cận, sơ lược về nội dung thực hiện cũng như tổng quan bố cục báo cáo.

Chương 2: PHÂN TÍCH XÁC ĐỊNH BÀI TOÁN: Tìm hiểu thực trạng, khả năng phát triển, đối thủ cạnh tranh, từ đó đưa ra chức năng mà ứng dụng cần đáp ứng.

Chương 3: TÌM HIỂU VÀ LỰA CHỌN CÔNG NGHỆ CHO ỨNG DỤNG: Tổng

quan các công nghệ trong ứng dụng.

Chương 4: PHÂN TÍCH VÀ THIẾT KẾ HỆ THỐNG: Dựa trên công nghệ đã lựa chọn, phân tích và thiết kế hệ thống dựa trên các tầng: Thiết kế dữ liệu, thiết kế kiến trúc và thiết kế giao diện.

Chương 5: KẾT LUẬN VÀ HƯỚNG PHÁT TRIỂN: Trình bày kết luận và hướng phát triển cho ứng dụng.

# CHƯƠNG 2. PHÂN TÍCH XÁC ĐINH BÀI TOÁN

## Khảo sát thực trạng

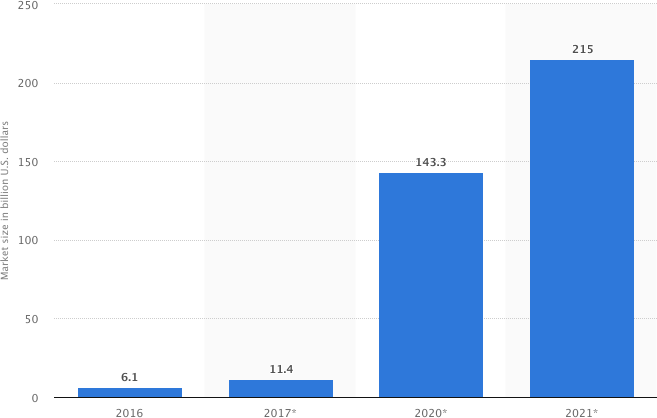
Nhận dạng ký tự quang học (tiếng Anh: Optical Character Recognition, viết tắt là OCR), là loại phần mềm máy tính được tạo ra để chuyển các hình ảnh của chữ viết tay hoặc chữ đánh máy (thường được quét bằng máy scanner) thành các văn bản tài liệu. OCR được hình thành từ một lĩnh vực nghiên cứu về nhận dạng mẫu, trí tuệ nhận tạo và machine vision. Mặc dù công việc nghiên cứu học thuật vẫn tiếp tục, một phần công việc của OCR đã chuyển sang ứng dụng trong thực tế với các kỹ thuật đã được chứng minh. Nhận dạng ký tự quang học (dùng các kỹ thuật quang học chẳng hạn như gương và ống kính) và nhận dạng ký tự số (sử dụng máy quét và các thuật toán máy tính) lúc đầu được xem xét như hai lĩnh vực khác nhau. Bởi vì chỉ có rất ít các ứng dụng tồn tại với các kỹ thuật quang học thực sự, bởi vậy thuật ngữ Nhận dạng ký tự quang học được mở rộng và bao gồm luôn ý nghĩa nhận dạng ký tự số. Đầu tiên hệ thống nhận dạng yêu cầu phải được huấn luyện với các mẫu của các ký tự cụ thể. Các hệ thống "thông minh" với độ chính xác nhận dạng cao đối với hầu hết các phông chữ hiện nay đã trở nên phổ biến. Một số hệ thống còn có khả năng tái tạo lại các định dạng của tài liệu gần giống với bản gốc bao gồm: hình ảnh, các cột, bảng biểu, các thành phần không phải là văn bản.

Hiện nay, công nghệ AR có thể đem đến những trải nghiệm cho người dùng hết sức độc đáo, chưa từng có tiền lệ. Có thể nhận định, tiềm năng phát triển của AR trong tương lai rất sáng sủa. Tuy nhiên, tồn tại một số luồng ý kiến cho rằng: AR rốt cuộc cũng chỉ là một chiêu thức tiếp thị hào nhoáng của các nhà sản xuất công nghệ. Ý kiến này không phải là không đúng, vì các ứng dụng AR tại thời điểm hiện tại chưa mang tính thực tiễn cao, và chưa áp dụng được trên phạm vi rộng lớn. Do đó, để AR trở thành công nghệ của tương lai, các nhà sản xuất cần tập trung nguồn lực vào cả hai mặt: nội dung và phần cứng. Trên thực tế, điện thoại và máy tính bảng sẽ không chỉ là địa điểm duy nhất để thực tế tăng cường phát triển. Mục tiêu cuối cùng của thực tế tăng cường là tạo ra sự

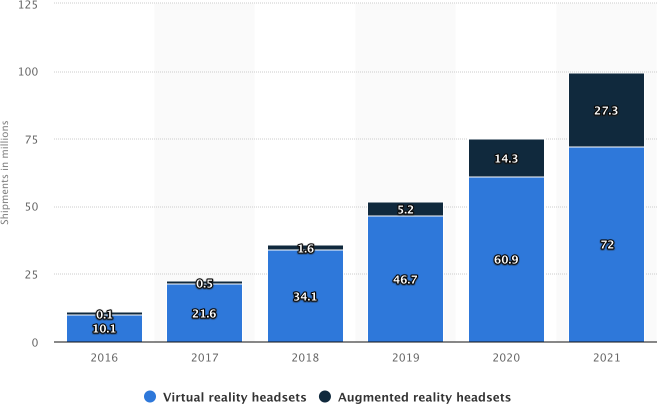
thuận tiện và tự nhiên, do đó, tương lai điện thoại và máy tính bảng sẽ được thay thế. Ngay cả kính cũng có thể mang hình dạng mới, các nhà khoa học đang nghiên cứu một loại kính thông minh dành cho người mù.

Dưới đây là những số liệu thú vị thể hiện tiềm năng của thị trường AR và VR. Mình nghĩ những số này chưa phản ánh đúng tiềm năng của thị trường này, đặc biệt khi mà các ông lớn đã nhảy vào thị trường.

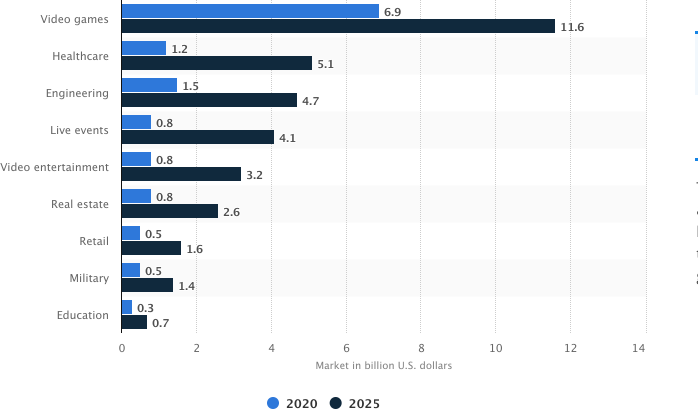
Nguồn: [www.statista.com](http://www.statista.com/)



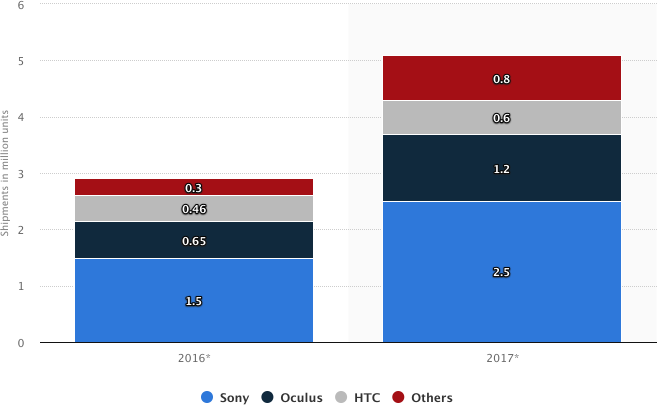
*Hình 2.1 - Độ lớn thị trường của AR và VR trong vài năm tới*



*Hình 2.2 - Số lượng thiết bị bán ra trong vài năm tới*



*Hình 2.3 - Độ lớn thị trường của các thiết bị AR và VR vào năm 2020 và 2025*



*Hình 2.4 - Số thiết bị bán ra theo từng hãng*

Theo báo cáo của [quý 3 năm 2017](https://www.canalys.com/newsroom/media-alert-virtual-reality-headset-shipments-top-1-million-first-time) thì:

* Sony bán được 490,000 thiết bị PlayStation VR (giá $399)
* Oculus bán được 210,000 thiết bị Rift (giá $399)
* HTC bán được 160,000 thiết bị Vive (giá $599)

Nghiên cứu này chỉ ra rằng:công nghệ AR/VR đang phát triển vượt trội.Bạn sẽ không tưởng tượng được công nghệ thay đổi chóng mặt như thế nào đâu, cho nên đừng ngạc nhiên nếu 1–2 năm nữa sẽ có các kính VR/AR không to tổ bố như hiện tại, chỉ to như cái kính cận thôi và có thể nghe gọi bình thường, dùng như một chiếc điện thoại và còn hơn thế nữa với mức giá tầm mà ai cũng có thể mua được.

## Các hệ thống liên quan

Nhận thấy với các mục tiêu mà ứng dụng cho khách hàng đã đề ra như trên, nhóm đã tiến hành tìm hiểu và khảo sát một số hệ thống đã có sẵn có mục tiêu tương tự, cũng như các đối thủ cạnh tranh để tìm ra ưu và khuyết điểm của từng hệ thống.

## AR animal



*Hình 2.5 - Ứng dụng AR animal*

### *Mô tả*

AR animal là ứng dụng giáo dục mầm non về động vật. Nó bao gồm 5 loại thẻ nhận thức: thẻ động vật trên cạn, thẻ sinh vật biển, thẻ chim, thẻ côn trùng và thẻ khủng long. Dùng máy ảnh để ghi lại hình động vật trên thẻ, mô hình 3D sẽ xuất hiện từ thẻ, giống như chúng đang sống trong điện thoại hoặc máy tính bảng. Sản phẩm được dành riêng để giúp trẻ học thông qua trải nghiệm tương tác thực tế. Đồng thời, nó rất hữu ích để phát triển trí thông minh, trau dồi trí tưởng tượng và thúc đẩy mối quan hệ cha mẹ và con cái.

### *Ưu điểm*

* Bao gồm năm loạt động vật, tổng số 100 động vật.
* Động vật 3D sinh động.
* Có thể tương tác với động vật.

### *Khuyết điểm*

* Sử dụng thẻ để nhận dạng động vật.
* Giá thành thẻ khá cao.

## AR Animals



*Hình 2.6 - Ứng dụng AR Animals*

### *Mô tả*

AR Animals là ứng dụng thực tế ảo dành cho di động giúp người dùng hiển thị các loài động vật dưới dạng hình ảnh 4D chân thực và sống động, chỉ cần dùng điện thoại cài [AR Animals](https://taimienphi.vn/download-ar-animals-79487) quét các thẻ bài in hình 4D riêng lẻ là có thể quan sát nhiều loại động vật khác nhau đang cử động. AR Animals này còn cung cấp chế độ xem với chuyển động xoay hình các con vật theo bất kỳ hướng nào để thuận tiện nhất cho việc quan sát.

### *Ưu điểm*

* Dễ sử dụng.
* Động vật 3D sinh động.
* Có thể tương tác với động vật.

### *Khuyết điểm*

* Sử dụng thẻ để nhận dạng động vật
* Hay bị lỗi không vào được.
* Rất ít động vật.

## Planet AR - Animals



*Hình 2.7 - Ứng dụng Plant AR - Animals*

### *Mô tả*

Planet AR - Động vật là một ứng dụng phần mềm miễn phí từ danh mục Giải trí. Ứng dụng với thẻ hiện động vật trong chế độ 3D, cho phép bạn quay, xoay và phóng to và thu nhỏ từng con vật.

### *Ưu điểm*

* Dễ sử dụng.
* Động vật 3D sinh động.

### *Khuyết điểm*

* Sử dụng thẻ để nhận dạng động vật
* Hay bị lỗi không vào được.
* Rất ít động vật.

## Tổng hợp

Sau khi phần tích một số hệ thống liên quan, nhóm nhận thấy có thể áp dụng một số tính năng đã có sẵn và khắc phục một số khuyết điểm như sau

* *Ưu điểm cần tiếp thu*
  + Cung cấp đa dạng các loài động vật.
  + Mô hình AR sinh động
  + Có thể tương tác với mô hình.
* *Khuyết điểm*
  + Phải sử dụng thẻ để nhận dạng động vật.
  + Mô hình vật động chỉ xuất hiện trên những tấm thẻ làm cho trải nghiệm ít đi phần thú vị.

Dựa trên khuyết điểm đó nhóm muốn tạo ra ứng dụng có thể cho phép người dùng trải nghiệm công nghệ thực tế ảo tăng cường. Người dùng không phải mua những tấm thẻ để có thể sử dụng được ứng. Người dùng có thể tùy chọn con vật, tùy chọn mặt phẳng mà họ thích để đặt chúng lên.

## Yêu cầu chức năng

Ứng dụng sẽ bao gồm các chức năng cho người dùng như sau:

* Nhận diện chữ chính xác.
* Xử lý với thời gian thực.
* Hỗ trợ đa ngôn ngữ.
* Biểu diễn mô hình thực tế ảo tăng cường.

## Yêu cầu phi chức năng

* Giao diện người dùng: Ứng dụng sẽ duy trì giao diện thân thiện, dễ sử dụng, hoạt động tốt trên nền tảng Android.
* Khả năng mở rộng: Ứng dụng có thể mở rộng dựa trên lượng động vật sử dụng và số lượng ngôn ngữ hỗ trợ.
* Độ chính xác: Ứng dụng cần đảm bảo độ chính xác của từ dạng chữ.
* Tính trực quan: Ứng dụng cung cấp mô hình động vật sinh động.
* Xử lí ngoại lệ: Các xử lí ngoại lệ phải được hiển thị cho người dùng nếu chúng xảy ra.

# CHƯƠNG 3. TÌM HIỂU VÀ LỰA CHỌN CÔNG NGHỆ

## Tổng quan về các công nghệ trong ứng dụng

## Tesseract OCR

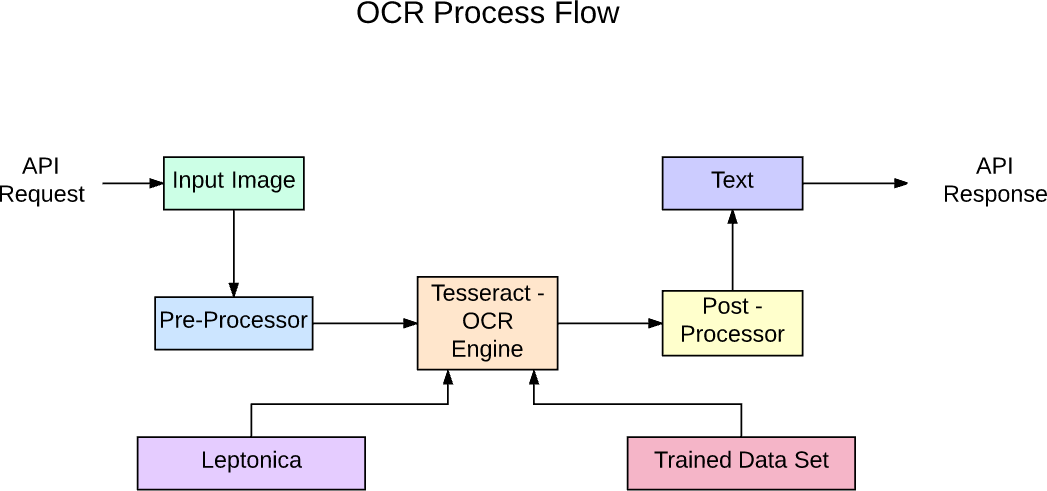
* + - 1. **Giới thiệu**



*Hình 3.1 - Tesseract OCR*

Nhận dạng ký tự quang học (tên tiếng anh là Optical Character Recognition – OCR) là kỹ thuật được sử dụng để chuyển đổi ảnh văn bản sang dạng văn bản có thể chỉnh sửa trong máy tính. Đầu vào của quá trình này là tập tin hình ảnh và đầu ra sẽ là các tập tin văn bản chứa nội dung là các chữ viết, ký hiệu có trong hình ảnh đó.

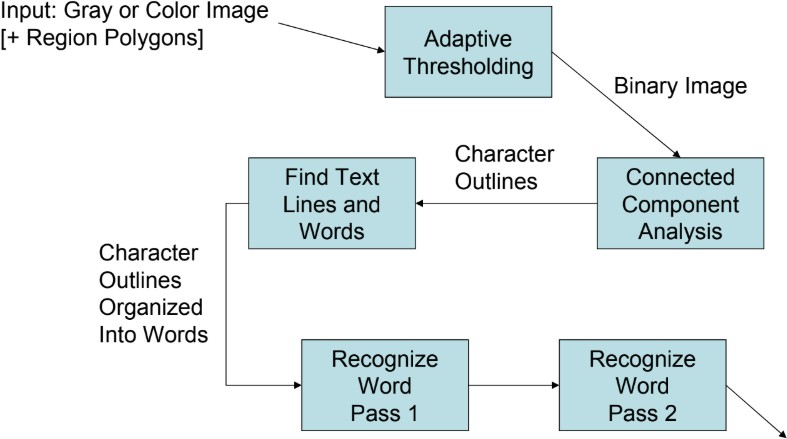
Tesseract là một hệ thống nhận dạng ký tự quang học (OCR). Nó được sử dụng để chuyển đổi các tài liệu hình ảnh thành các tài liệu PDF hoặc Word có thể chỉnh sửa / tìm kiếm được. Đây là một phần mềm mã nguồn mở miễn phí chạy qua Giao diện dòng lệnh (CLI). Tesseract được coi là một trong những công cụ OCR mã nguồn mở chính xác nhất hiện có và sự phát triển của nó đã được Google tài trợ từ năm 2006. Điều đó có thể nói, khả năng của nó có thể bị hạn chế hơn so với phần mềm thương mại như Adobe Acrobat Pro và ABBYY FineReader. Tuy nhiên, vì là phần mềm nguồn mở, bất kỳ ai có kiến thức lập trình đều có thể chỉnh sửa mã phía sau Tesseract và giúp nó tìm hiểu những gì bạn cần làm. Nó có thể được sử dụng trên các máy Mac, Windows và Linux.



*Hình 3.2 - Flow cơ bản của Tesseract*

Tesseract ban đầu được thiết kế để nhận dạng các từ tiếng Anh trên ngôn ngữ hệ Latinh. Sau này, nhờ sự cố gắng của nhiều nhà phát triển mà các phiên bản của Tesseract đã có thể nhận diện các ngôn ngữ khác ngoài hệ Latinh như tiếng Trung, tiếng Nhật và tương thích với các ký tự trong bảng mã UTF-8. Việc nhận dạng các ngôn ngữ mới trên Tesseract có thể thực hiện được nhờ vào việc huấn luyện dữ liệu. Từ phiên bản 3.0 trở đi, Tesseract đã có thể hỗ trợ thêm nhiều dạng ngôn ngữ mới và mở rộng thêm việc huấn luyện theo font chữ. Bởi vì ban đầu, bộ Tesseract được huấn luyện để nhận diện từ chính xác nhất trên một số loại font mặc định, nếu sử dụng các font chữ khác để nhận diện thì có thể kết quả sẽ không có độ chính xác cao khi làm việc với các loại font được cài đặt sẵn trong dữ liệu huấn luyện.

## Các thức hoạt động





*Hình 3.3 - Cơ chế hoạt động của Tesseract*

Trước tiên, hình ảnh sẽ được phân tích để tìm ra các vùng kết nối (connected component). Bước này cho phép OCR dễ dàng nhận biết những vùng ký tự ngược để có thể nhận diện những ký tự bên trong. Trong Tesseract, những vùng chứa ký tự này được gọi là Blob.

Tiếp đến, những blob này sẽ tiếp tục được phân tích để tìm ra các dòng, rồi đến các ký tự. Việc tìm các dòng sẽ được xử lý bởi thuật toán dựa vào vùng ký tự, cỡ chữ cùng

toạ độ (trục x). Trong quá trình này, các blob cũng có thể được ghép với nhau nếu **OCR** nhận thấy chúng chứa các ký tự trong cùng một dòng. Những **blob** được ghép phải trùng ít nhất 50% theo chiều ngang. Sau đó, các đường cơ sở (baseline) cũng được tìm kiếm nhờ vào việc quét các dòng đã được xác định.

Sau khi đã xác định được các dòng ký tự cùng các đối số tương ứng, dòng ký tự sẽ được chia nhỏ thành các từ dựa vào các ký tự phân cách. Lúc này, văn bản cố định sẽ được chia nhỏ và tiến hành nhận diện. Trong khi đó, văn bản không cố định hoặc chưa chắc chắn thì sẽ được chia nhỏ thành các từ dù chưa chắc chắn. Nhưng nhờ vào bước nhận diện, chúng ta sẽ thu được kết quả cuối cùng chính xác hơn.

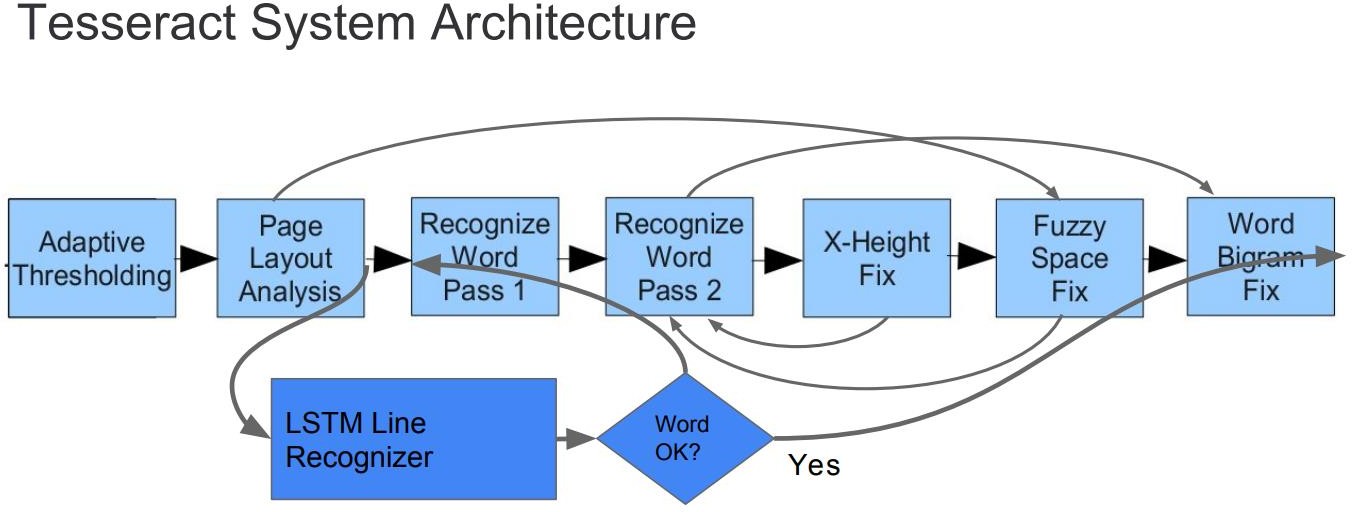


*Hình 3.4 - Các khối blob sau khi tách*

Bước vào quá trình nhận diện, input của chúng ta sẽ được đánh giá, phân tích hai lần. Ở lần đầu tiên, OCR sẽ nhận diện ký tự với kết quả phân tích ở bước trước đó. Các kết quả nhận diện thoả mãn yêu cầu sẽ được đưa vào tập tin huấn luyện để hỗ trợ cho quá trình nhận diện lần thứ hai với các kết quả chưa đạt yêu cầu. Đương nhiên, việc xác nhận kết quả có thoả mãn yêu cầu hay không cần phải dựa trên nhiều tiêu chí vì nhận diện nội dung phải trải qua một quá trình lặp đi lặp lại gồm các bước nhận diện ký tự, ghép ký tự và so khớp với từ điển. Các tiêu chí đó bao gồm khoảng cách của các ký tự, độ phù hợp với từ điển và khoảng cách đến các dấu câu.

Cuối cùng, OCR sẽ xử lý những dấu cách không rõ ràng cùng với xem xét các giả thiết khác cho việc định vị những ký tự in hoa nhỏ để đi đến kết quả cuối cùng.

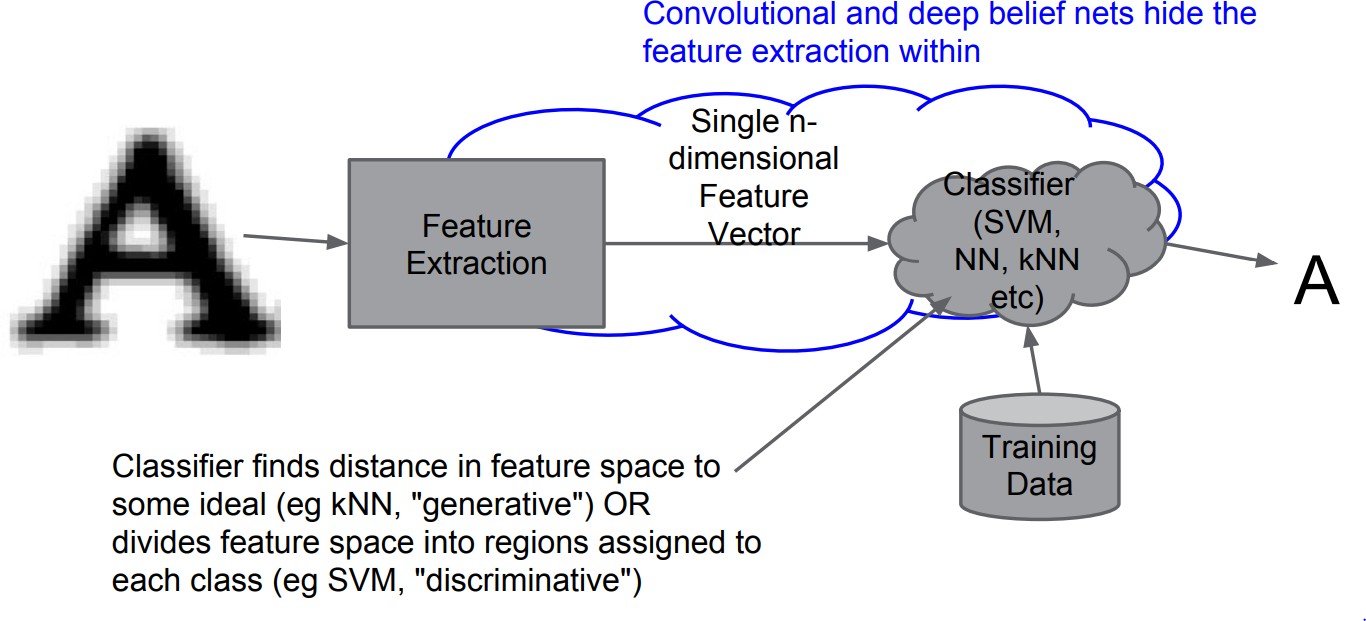
## Cấu trúc hệ thống



*Hình 3.5 - Cấu trúc của Tesseract*

## Đặc điểm phân loại

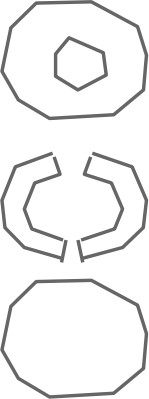
### *Nhận dạng kí tự*



*Hình 3.6 - Cơ chế nhận dạng kí tự*

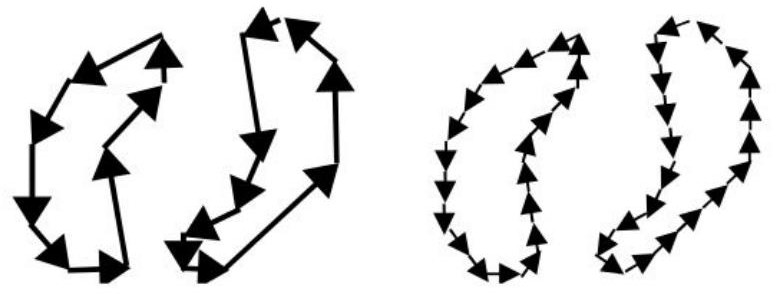
### *Nhận dạng cấu trúc*

Các khả năng chữ “O”:

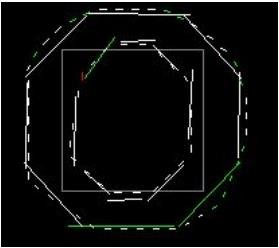
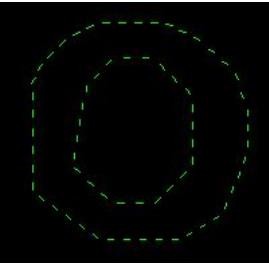
* Tiêu chuẩn
* Hỏng
* Đầy

*Hình 3.7 - Các khả năng của chữ “O”*

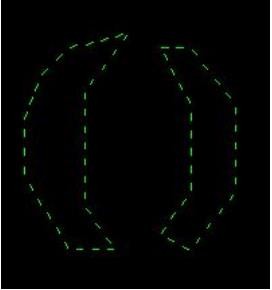
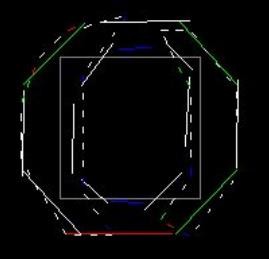
Các phân đoạn đa giác Các phân đoạn nhỏ hơn



*Hình 3.8 - Các phân đoạn đoạn chữ “O”*



*Hình 3.9 - Cơ chế chuyển đổi nhận dạng*

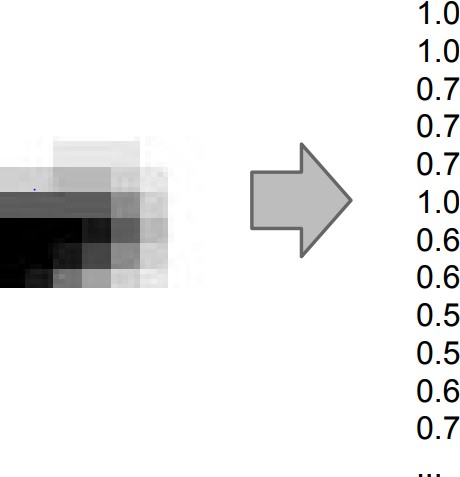
*Hình 3.10 - Cơ chế chuyển đổi nhận dạng*

### *Deep Belief nets*

1 pixel = 1 đặc tính kích thước

1 ký tự (ví dụ 32x32) = ~ 1K vectơ đặc tính kích thước

* Các đặc tính được học
* Thường là các cạnh
* Sự phụ thuộc đặc điểm giữa các pixel phải cũng được học



*Hình 3.11 - Cơ chế Deep Belief nets*

### *Biểu đồ độ dốc*

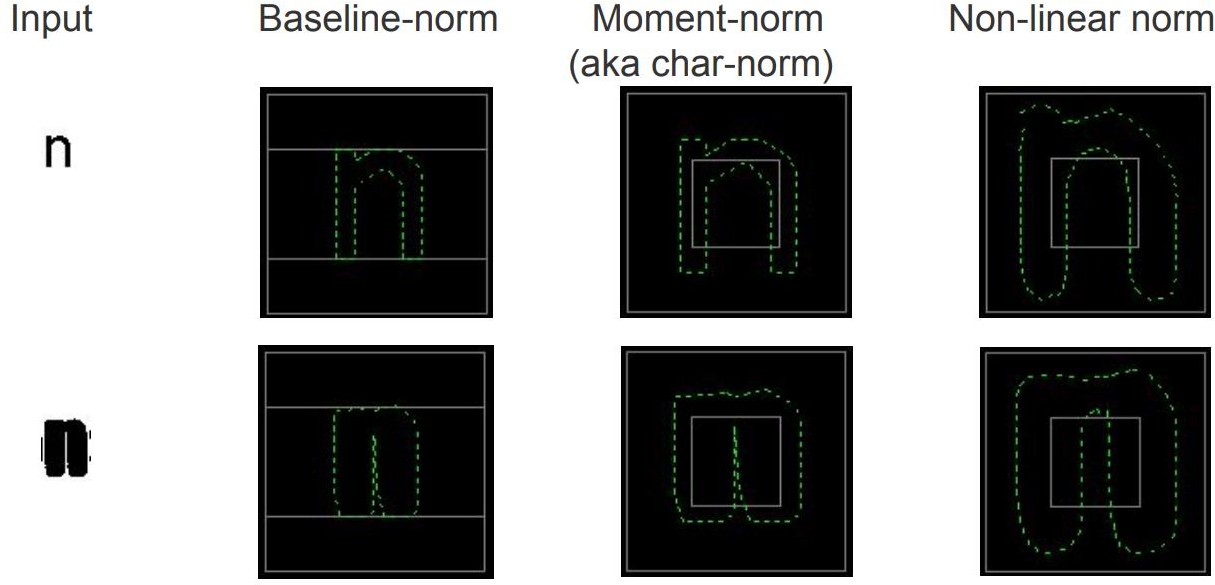
* Xác định vùng ký tự
* Tính toán độ dốc
* Biểu đồ độ dốc ánh xạ tới vectơ đặc tính



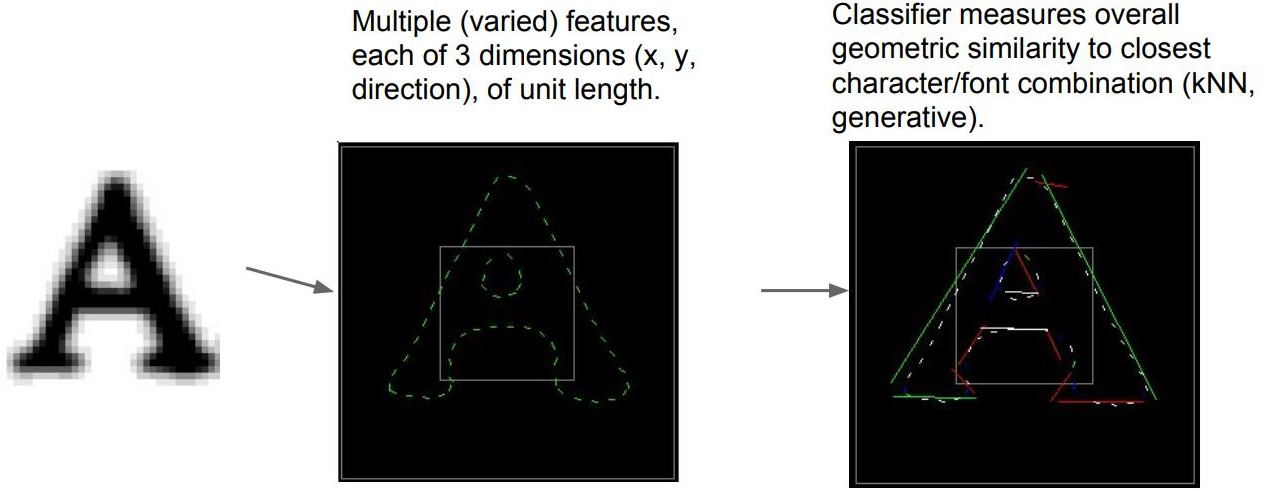
*Hình 3.12 - Cơ chế độ dốc*

### *Chuẩn hóa phân loại*

* Định mức cơ sở
* Định mức ký tự
* Đinh mức phi tuyến tính



*Hình 3.13 - Cơ chế chuẩn hóa phân loại*



*Hình 3.14 - Cơ chế chuẩn hóa phân loại*

## Long Short Term Memory (LSTM)

Tesseract có sự trưởng thành với phiên bản 3.x khi nó bắt đầu hỗ trợ nhiều định dạng hình ảnh và dần dần thêm một số lượng lớn các tập lệnh (ngôn ngữ). Tesseract 3.x dựa trên các thuật toán thị giác máy tính truyền thống. Trong vài năm qua, các phương pháp dựa trên Deep Learning đã vượt qua các kỹ thuật máy học truyền thống bằng một

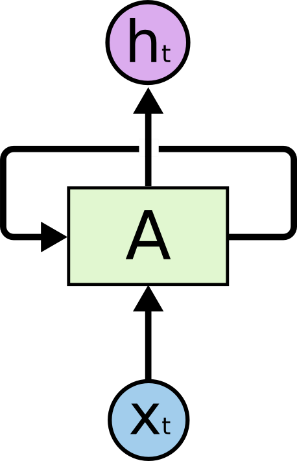
lợi thế rất lớn về độ chính xác trong nhiều lĩnh vực Tầm nhìn Máy tính. Nhận dạng chữ viết tay là một trong những ví dụ nổi bật. Vì vậy, đó chỉ là vấn đề thời gian trước khi Tesseract cũng có một công cụ nhận dạng dựa trên Deep Learning.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Engine** | **Total char errors** | **Word Recall Errors** | **Word Precision Errors** | **Walltime** | **CPUtime\*** |
| Tess 3.04 | 13.9 | 30 | 31.2 | 3.0 | 2.8 |
| Cube | 15.1 | 29.5 | 30.7 | 3.4 | 3.1 |
| Tess+ Cube | 11.0 | 24.2 | 25.4 | 5.7 | 5.3 |
| LSTM | 7.6 | 20.9 | 20.8 | 1.5 | 2.5 |

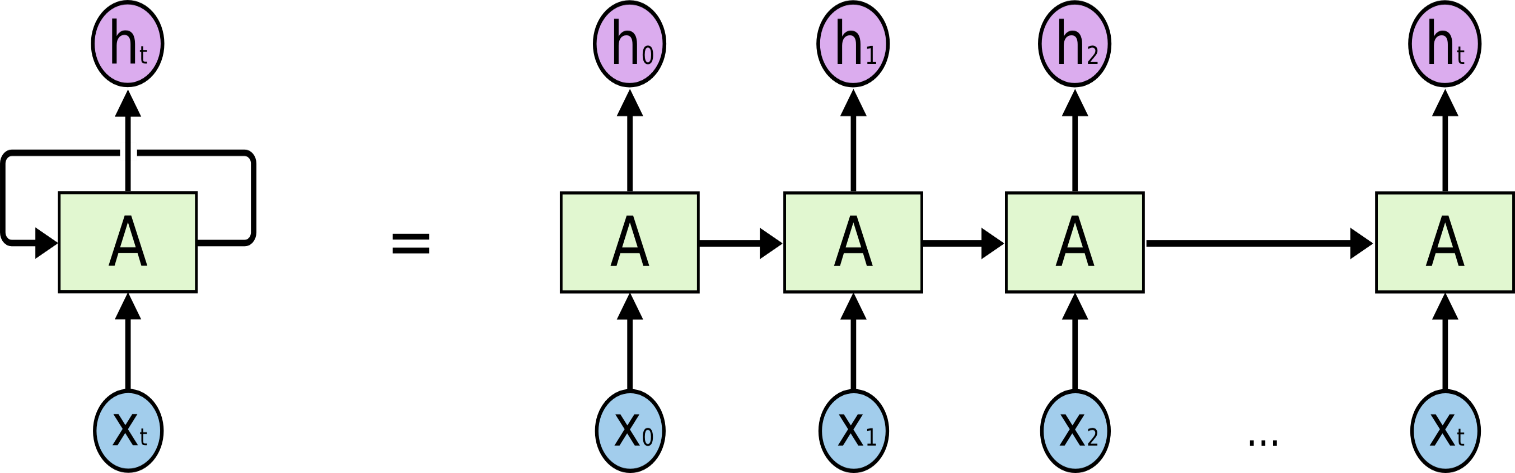
*Bảng 3.1 - Bảng so sánh các công cụ nhận dạng chữ*

Trong phiên bản 4, **Tesseract** đã triển khai công cụ nhận dạng dựa trên Bộ nhớ ngắn hạn (**LSTM**). **LSTM** là một loại Mạng thần kinh tái phát (**RNN**).

Mạng nơ-ron hồi quy **RNN** (Recurrent Neural Network) chứa các vòng lặp bên trong cho phép thông tin có thể lưu lại được.

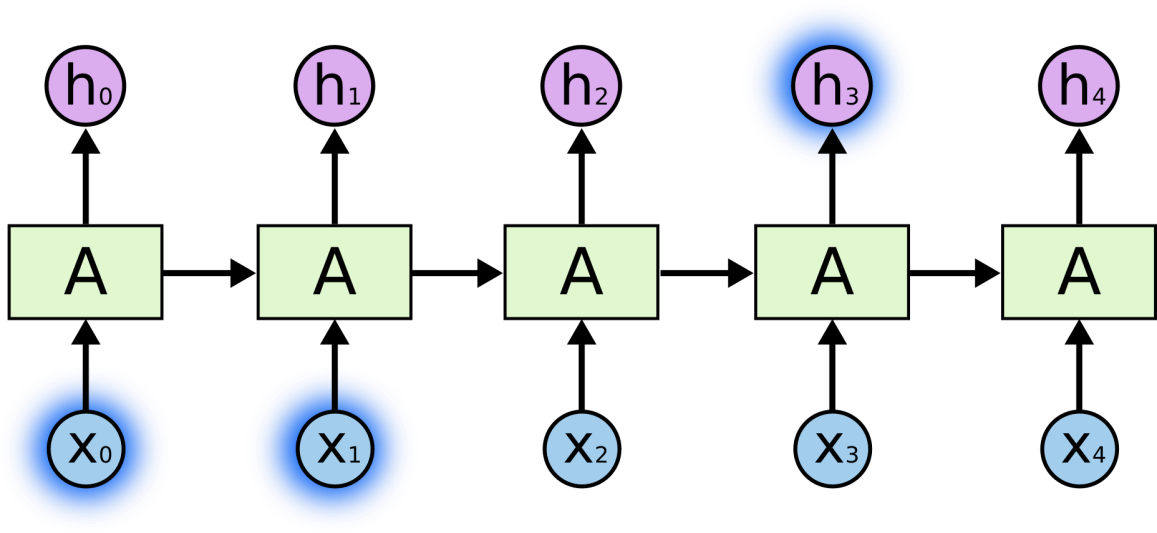


*Hình 3.15 - Mạng nơ-ron hồi quy*

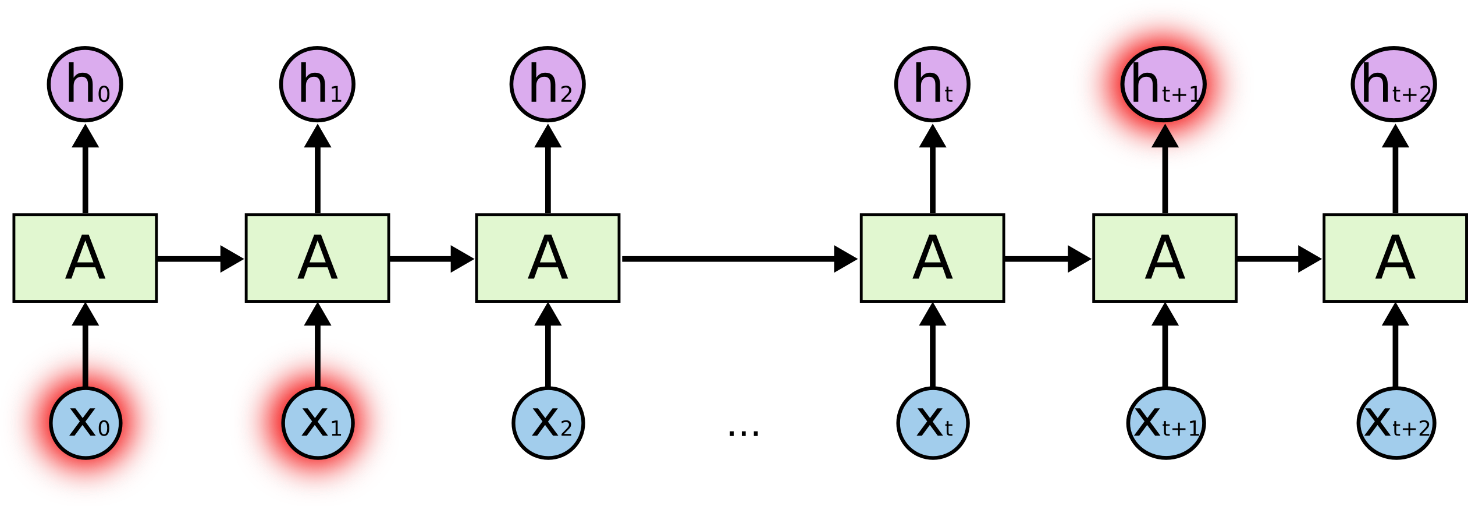
Một mạng nơ-ron hồi quy có thể được coi là nhiều bản sao chép của cùng một mạng, trong đó mỗi đầu ra của mạng này là đầu vào của một mạng sao chép khác.

*Hình 3.16 - Mạng nơ-ron hồi quy*

Ví dụ, ta có câu: “*The clouds in the sky*” thì ta chỉ cần đọc tới “*The clouds in the*” là đủ biết được chữ tiếp theo là “*sky*” rồi. Trong tình huống này, khoảng cách tới thông tin có được cần để dự đoán là nhỏ, nên **RNN** hoàn toàn có thể học được.



*Hình 3.17 - Mạng nơ-ron hồi quy*

Nhưng trong nhiều tình huống ta buộc phải sử dụng nhiều ngữ cảnh hơn để suy luận. Ví dụ, dự đoán chữ cuối cùng trong đoạn: “*I grew up in France… I speak fluent French.*”. Rõ ràng là các thông tin gần (”*I speak fluent*”) chỉ có phép ta biết được đằng sau nó sẽ là tên của một ngôn ngữ nào đó, còn không thể nào biết được đó là tiếng gì. Muốn biết là tiếng gì, thì ta cần phải có thêm ngữ cảnh “*I grew up in France*” nữa mới có thể suy luận được.

*Hình 3.18 - Mạng nơ-ron hồi quy*

Mạng bộ nhớ dài-ngắn (**Long Short Term Memory networks**), thường được gọi là

**LSTM** - là một dạng đặc biệt của **RNN**, nó có khả năng học được các phụ thuộc xa. **LSTM**

được giới thiệu bởi Hochreiter & Schmidhuber (1997), và sau đó đã được cải tiến và phổ biến bởi rất nhiều người trong ngành.

**LSTM** được thiết kế để tránh được vấn đề phụ thuộc xa (long-term dependency). Việc nhớ thông tin trong suốt thời gian dài là đặc tính mặc định của chúng, chứ ta không cần phải huấn luyện nó để có thể nhớ được. Tức là ngay nội tại của nó đã có thể ghi nhớ được mà không cần bất kì can thiệp nào.

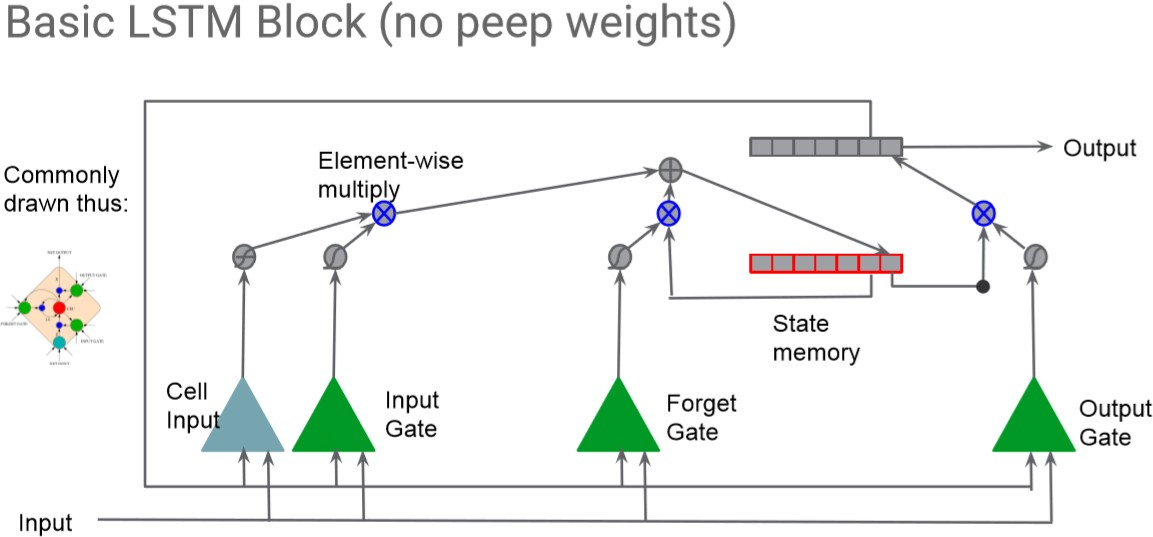
## Đánh giá Tesseract

### *Code cleanup*

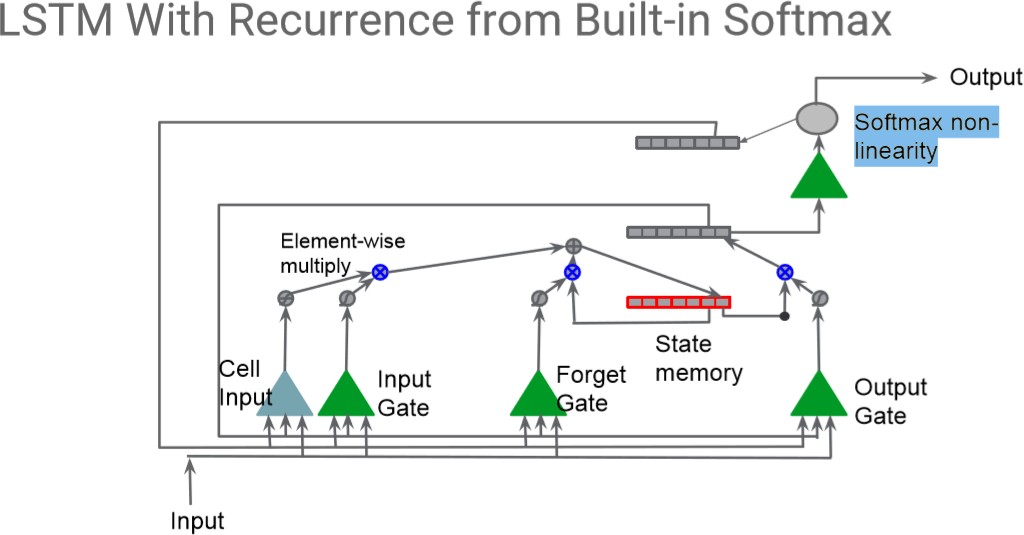
* Đã hoàn tất chuyển đổi sang C ++
* Khả năng tương thích luồng: có thể chạy nhiều trường hợp trong các luồng riêng biệt
* Khả năng đa ngôn ngữ (đơn và hỗn hợp): đối phó với jpn, heb, hin và hỗn hợp như hin + eng
* Đã xóa nhiều giới hạn mã hóa cứng , ví dụ: trên bộ ký tự, dawss
* Tìm kiếm chùm tia mới
* Bộ kết quả để truy cập dữ liệu nội bộ
* API phân loại chung hơn cho các thử nghiệm plug-n-play
* Đã xóa rất nhiều mã chết, bao gồm cả lớp IMAGE.

### *Hiện đại hóa thuật toán*

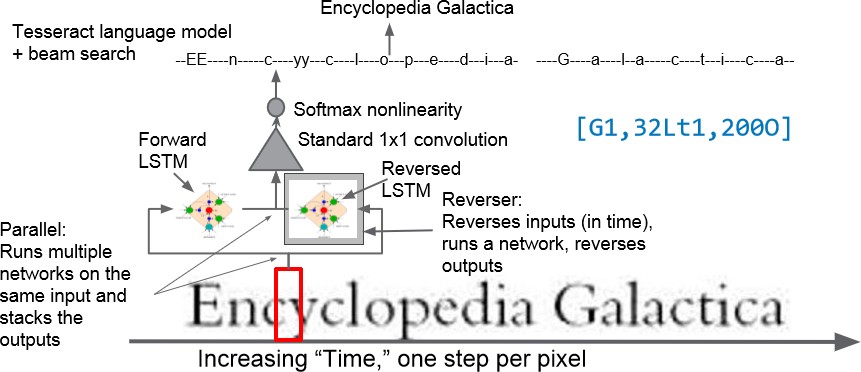
* LSTM dựa trên triển khai Python của OCROpus.
* Các khả năng mở rộng bao gồm 2-D, kích thước đầu vào biến đổi.
* Tích hợp hoàn toàn với Tesseract ở cấp độ từ tương tự nhóm.
* Trực quan hóa với Viewer API .
* Bao gồm mã traning.
* Song song với openMP.



*Hình 3.19 - LSTM cơ bản*



*Hình 3.20 - LSTM với softmax*



*Hình 3.21 - Tesseract sử dụng LSTM*

## Vì sao chọn Tesseract?

Do tính chất của bộ dữ liệu đào tạo Tesseract, việc nhận dạng ký tự kỹ thuật số được ưa thích hơn, mặc dù Tesseract OCR cũng có thể được sử dụng để nhận dạng chữ viết tay. Tesseract OCR với phiên bản ổn định 4.0.0 mới nhất dựa trên LSTM (bộ nhớ ngắn hạn dài) nhận diện nhanh, chính xác hơn.

Phiên bản chính thức của Tesseract OCR cho phép các nhà phát triển xây dựng ứng dụng của riêng họ bằng API C hoặc C ++. Theo thời gian, cộng đồng đã tạo ra các phiên bản riêng của họ về các công cụ bên ngoài, trình bao bọc và thậm chí các dự án đào tạo. Lý do chọn Tesseract vì:

* Nó đơn giản và dễ sử dụng
* Nó hỗ trợ phiên bản 4.0.0 (tại thời điểm viết bài này)
* Các ý kiến và giải thích trong tập tin rất chi tiết

## AR – AR Core

* + - 1. **AR là gì?**



*Hình 3.22 - Công nghê AR*

Augmented Reality (AR) là một công nghệ gồm thế giới xung quanh và thêm nội dung ảo lên sao cho giống như nó thực sự tồn tại trong thế giới thực. AR là một trải nghiệm tương tác của môi trường trong thế giới thực, nơi các vật thể sống trong thế giới thực được "tăng cường" bởi thông tin nhận thức do thiết bị tạo ra. Công nghệ thay đổi nhận thức liên tục của một người về môi trường thế giới thực, trong khi thực tế ảo thay thế hoàn toàn môi trường thế giới thực của người dùng bằng mô phỏng.

AR là một cách chồng các vật thể hoạt hình ảo lên môi trường xung quanh trong thế giới thực bằng thiết bị thông minh. Một môi trường sống được tăng cường bởi dữ liệu kỹ thuật số - âm thanh, hình ảnh, video. Có 2 nhóm ứng dụng AR:

* Các ứng dụng dựa trên điểm đánh dấu.
* Ứng dụng dựa trên vị trí.

Những ứng dụng đầu tiên hoạt động nhận dạng hình ảnh bằng camera để quét hình ảnh (điểm đánh dấu) và sau đó thêm hình ảnh ảo trên màn hình điện thoại. Ví dụ: nhiều ứng dụng đọc mã QR và trình bày thông tin bổ sung. Các ứng dụng AR dựa trên vị trí sử dụng GPS để định vị các địa điểm gần đó và / hoặc để cung cấp chỉ đường, v.v.

## AR trong thực tiễn

Công nghệ luôn cách mạng hóa cách chúng ta làm việc và sống; xem xét máy tính, điện thoại, máy giặt và nhiều hơn nữa. Chúng cho phép chúng ta có cuộc sống dễ dàng và hiệu quả hơn, và thực tế tăng cường (AR) là bước tiếp theo trong hành trình đó. Các doanh nghiệp ngày nay đang áp dụng AR:

### *Bán lẻ*

Các nhà bán lẻ đang ưu tiên trải nghiệm khách hàng và kỹ thuật số. Với sự mờ nhạt giữa ranh giới giữa mua sắm kỹ thuật số và vật lý, AR sẽ cho phép các doanh nghiệp thu hẹp khoảng cách và giới thiệu những cách thức mới và cải tiến cho khách hàng mua sắm. Trong các cửa hàng quần áo kỹ thuật số có thể cho phép người dùng nhìn thấy mặt hàng quần áo ở định dạng 3D.

### *Ứng dụng công nghiệp*

Công nghệ AR trong công nghiệp có thể sẽ thay đổi số lượng công việc được thực hiện. Kỹ thuật viên trong lĩnh vực này sẽ có thể nhận được hỗ trợ trực tiếp từ nhân viên từ xa, người có thể chỉ ra các dấu hiệu, chỉ ra các vấn đề, mô hình chồng lên các vật phẩm như động cơ xe và những thứ tương tự, v.v. Nó loại bỏ sự phỏng đoán khỏi công việc thủ công và hợp lý hóa các quy trình.

### *Thiết kế & Sáng tạo*

Các ngành công nghiệp chú trọng không gian thiết kế và sáng tạo có thể sẽ là một số thị trường bị ảnh hưởng tích cực nhất bởi sự ra đời của thực tế tăng cường. Các công ty AR như Augment đã sử dụng công nghệ để giúp các doanh nghiệp. Phát triển các ứng

dụng cung cấp cho các công ty khả năng đặt các mô hình 3D vào không gian vật lý sẽ cho phép họ cung cấp tài liệu tiếp thị và bán hàng mạnh mẽ hơn. Các nghệ sĩ và người tạo mô hình CAD (thiết kế có sự trợ giúp của máy tính) cũng sẽ được hưởng lợi từ AR, vì nó mang lại cho họ phương tiện để kết xuất sản phẩm và các tài sản khác trong không gian 3D để trình bày và trình bày công việc dễ dàng hơn.

### *Đào tạo*

Làm thế nào để một công ty đào tạo nhân viên một cách hiệu quả và áp dụng cùng một khóa đào tạo trên các địa điểm, đặc biệt là nếu công ty nhượng quyền thương mại lớn? Còn những ứng dụng phức tạp, như trong lĩnh vực y tế thì sao? CAE chăm sóc sức khỏe đang tích hợp Microsoft Tủ HoloLens vào các chương trình đào tạo của mình, cho phép các bác sĩ thực hành các quy trình y tế phức tạp trong môi trường 3D. AR cho phép các công ty phát triển chương trình đào tạo phù hợp với từng nhân viên và cho phép nhân viên phát triển năng lực và sự tự tin về vai trò của họ.

### *Tương tác khách hàng*

Khán giả hiện đại thích ứng với công nghệ mới một cách nhanh chóng; 4.2 tỷ người đã được báo cáo là người dùng internet tích cực trong năm 2018 và đến năm 2020, cơ sở lắp đặt các thiết bị internet được dự báo sẽ tăng lên gần 31 tỷ. AR sẽ là bước tiếp theo, khi các công ty phát triển những cách mới để áp dụng nó. Một cách như vậy có thể được nhìn thấy với ứng dụng Magic Leap sườn Cheddar, cho phép các mạng tin tức tiêu đề Cheddarftime cung cấp nguồn cấp dữ liệu video trực tiếp vào không gian 3D cho người dùng, những người có thể đặt chúng ở bất cứ đâu trong môi trường xung quanh. Các doanh nghiệp áp dụng thực tế tăng cường cho các sản phẩm của họ có được sự tham gia trực quan trực tiếp với khán giả của họ.

## Giới thiệu AR Core

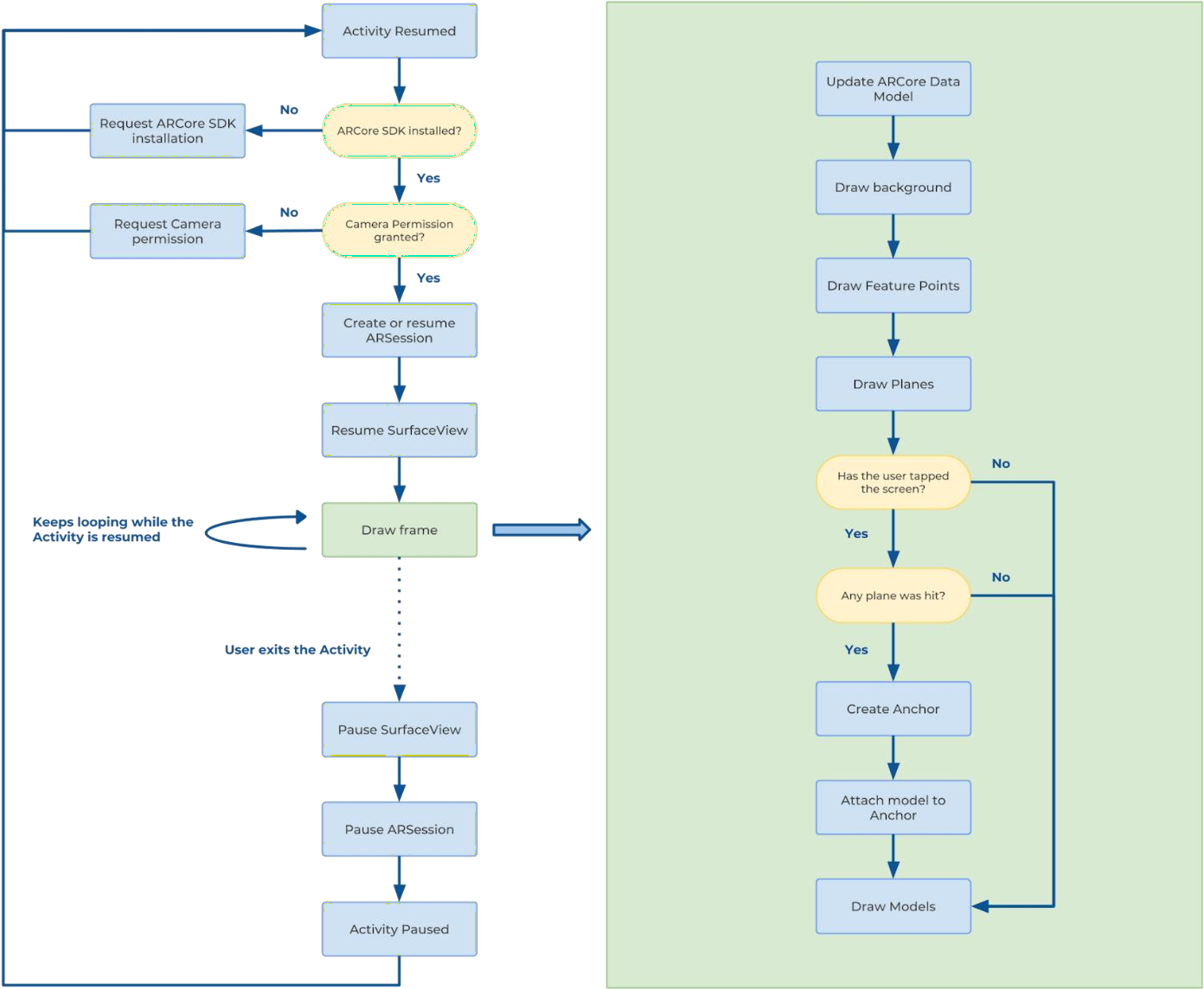


*Hình 3.23 - AR Core*

**ARCore** là một nền tảng của Google dùng để trải nghiệm thực tế ảo tăng cường. **ARCore** sử dụng các API khác nhau, giúp điện thoại của bạn có thể cảm nhận được môi trường xung quanh, hiểu được thế giới thực và tương tác với các thông tin trong thế giới thực. Một số API đã có sẵn trên Android và iOS để cho phép chia sẻ trải nghiệm về AR.

**ARCore** là một SDK linh hoạt, có thể làm việc với Java/OpenGL, Unity, Unreal Engine, và tập trung vào mục tiêu để mang lại trải nghiệm AR cho nhiều đối tượng người dùng hơn.

## Cách thức hoạt động



*Hình 3.24 - Cơ chế hoạt động ARCore*

Về cơ bản, ARCore sẽ làm 2 việc: theo dõi vị trí của thiết bị khi di chuyển và xây dựng hiểu biết sự hiểu biết riêng của nó về thế giới thực. Công nghệ motion tracking của ARCore sử dụng camera của điện thoại để xác định các điểm quan trọng và theo dõi các điểm đó di chuyển theo thời gian. Với sự kết hợp sự chuyển động của các điểm và đọc nó từ cảm biến quán tính của điện thoại, ARCore xác định cả vị trí hướng của điện thoại khi di chuyển qua các bề mặt. Ngoài việc xác định các điểm chính, ARCore có thể phát hiện các bề mặt phẳng ví dụ như mặt đất hoặc mặt bàn, và cũng có thể ước lượng được ánh

sáng của khu vực xung quanh nó. Kết hợp 2 khả năng này giúp ARCore xây dựng nên hiểu biết riêng của nó về thế giới xung quanh.

Sự hiểu biết của ARCore về thế giới thực giúp chúng ta có thể đặt các đối tượng, chú thích hoặc nhưng thông tin khác mà nó có thể tích hợp hoàn hảo với thế giới thực. Motion tracking có nghĩa là bạn có thể di chuyển xung quanh và nhìn thấy các đối tượng này ở bất kì góc độ nào, và thậm chí nếu bạn quay đi và rời khỏi phòng, khi quay lại, các đối tương, chú thích và thông tin vẫn ở nguyên vị trí mà bạn đặt nó.

## Các đặc điểm nổi bật

### *Theo dõi chuyển động*

Khi điện thoại của bạn di chuyển trong thế giới thực, ARCore sẽ đo đạc đồng thời ánh xạ, hoặc COM, để hiểu được rằng điện thoại có liên quan như thế nào với thế giới xung quanh nó. ARCore phát hiện các tính năng riêng biệt trực quan thông qua camera, gọi là feature points (điểm đặc trưng) và sử dụng các điểm này để tính toàn sự thay đổi về vị trí. Thông tin về thị giác được kết hợp với các phép đo quán tính từ IMU sẽ ước tính được pose (vị trí và hướng) của camera liên quan đến thế giới thực theo thời gian.

Bằng cách căn chỉnh tư thế của camera ảo sẽ hiển thị nội dung 3D của bạn với pose của camera được cung cấp bởi ARCore, lập trình viên sẽ có thể kết xuất các nội dung ảo từ một phối cảnh chính xác. Hình ảnh ảo được kết xuất có thể phủ lên trên của hình ảnh thu được từ camera, làm cho nó xuất hiện như thể nội dung ảo là một phần của thế giới thực.



*Hình 3.25 - Cơ chế theo dõi chuyển động*

### *Nhận biết được môi trường*

ARCore không ngừng nâng cao hiểu biết của nó về thế giới thực bằng cách phát hiện các điểm và mặt phẳng đặc trưng. ARCore tìm kiếm các cụm feature points xuất hiện nằm trên các bề mặt ngang hoặc dọc, ví dụ như bàn hay tường và làm các bề mặt này xuất hiện trong ứng dụng như là các mặt phẳng. ARCore cũng có thể xác định được ranh giới của các mặt phẳng và truyền các thông tin đó đến ứng dụng. Chúng ta có thể sử dụng thông tin đó để đặt các vật thể ảo trên bề mặt phẳng.

Vì ARCore sử dụng feature points để xác định mặt phẳng, mặt phẳng không có các kết cấu như tường trắng, có thể không được xác định chính xác.



*Hình 3.26 - Cơ chế nhận biết môi trường*

### *Ước lượng ánh sáng*

ARCore có thể phát hiện thông tin về ánh sáng của môi trường và hiệu chỉnh màu sắc của hình ảnh camera. Những thông tin này cho phép bạn chiếu sáng các đối tượng ảo của mình trong cùng điều kiện như môi trường xung quanh chúng, làm nó trở nên thực hơn.

*Hình 3.27 - Cơ chế ước lượng ánh sáng*

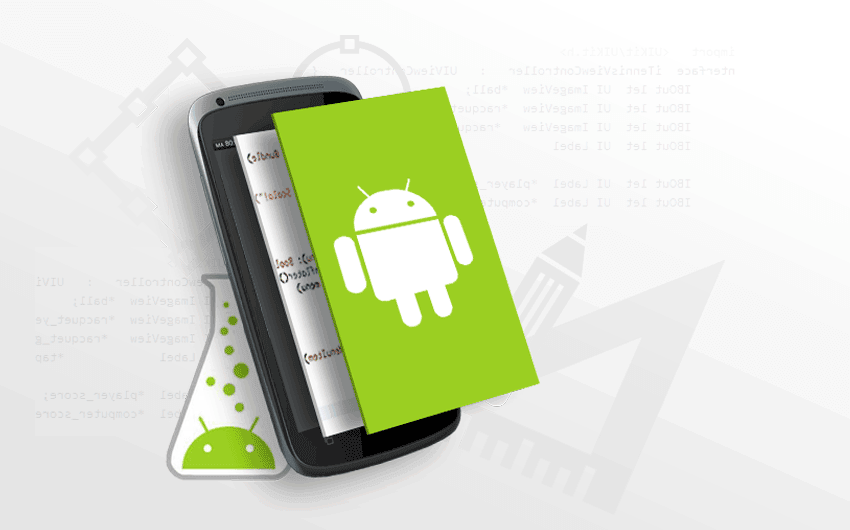
## Vì sao chọn AR Core?

Các ứng dụng ARCore có kích thước khá nhẹ. Một số ứng dụng hỗ trợ ARCore của Google như các ứng dụng giáo dục giúp người dùng có thể khám phá thế giới hoặc cơ thể con người, ngoài ra còn có một số game AR khác cũng hỗ trợ ARCore.

ARCore cũng có khả năng làm tăng trải nghiệm người dùng Android. Tất cả những gì cần cho sự thành công của ARCore là sự hỗ trợ của các developer, các công ty sử dụng thử nghiệm SDK và cung cấp các dịch vụ phát triển ứng dụng Android sáng tạo. Google tin tưởng vào việc cung cấp công nghệ này cho mọi người ngay cả trên web bằng cách xây dựng trình duyệt hỗ trợ nội dung AR và chạy trên cả thiết bị Android và iOS.

## Các công nghệ khác

* + - 1. **Android**



*Hình 3.28 - Hệ điều hành Android*

Android là một hệ điều hành có mã nguồn mở dựa trên nền tảng Linux được thiết kế dành riêng cho các thiết bị di động có màn hình cảm ứng như điện thoại thông minh và máy tính bảng. Ban đầu, hệ điều này này được phát triển bởi công ty Android, với sự

hỗ trợ tài chính từ Google, sau đó chính Google đã mua lại công ty này và tiếp tục phát triển Android trở thành một nền tảng hiệu quả hơn.

Là hệ điều hành mã nguồn mở kết hợp với việc được Google phát hành mã nguồn theo giấy phép Apache (một loại giấy phép ít bị ràng buộc), chính hai yếu tố này đã giúp cho các nhà phát triển thiết bị, các nhà mạng, các lập trình viên nhanh chóng tiếp cận, điều chỉnh và phân phối Android một cách tự do. Có thể nói hành động Google “cho không” Android là một bước tiến lớn trong việc biến đứa con này trở thành nền tảng điện thoại thông minh phổ biến nhất thế giới. Bằng chứng là về mặt ứng dụng, với bản chất “mở” của mình thì Android đã thu hút được một cộng đồng đông đảo các lập trình viên, chuyên viên phát triển ứng dụng, theo thống kê tháng 10 năm 2012 đã xuất hiện khoảng

* 1. ứng dụng Android, và số lượng tải từ Google Play, cửa hàng ứng dụng chính của Android được ước tính vào khoảng 25 tỷ lượt. Về mặt thiết bị thì nền tảng này liên tục được các nhà sản xuất thiết bị lựa chọn khi mà họ đang cần một hệ điều hành hội tụ những ưu điểm, vận hành nhẹ nhàng, có khả năng tinh chỉnh tốt, và tiết kiệm chi phí để chạy trên các thiết bị công nghệ cao của mình.

Các lý do chọn Android làm nền tảng phát triển ứng dụng:

### *Chi phí đầu tư thấp*

Android cung cấp bộ công cụ phát triển phần mềm (SDK) miễn phí cho cộng đồng nhà phát triển nhằm giảm thiểu chi phí phát triển và cấp phép. Phát triển ứng dụng trong Android, nó rẻ hơn so với các nền tảng khác. Tất cả các chi phí được dành cho việc thử nghiệm và triển khai ứng dụng. Điều đó có nghĩa là không phải đầu tư lớn vào phần đó.Đầu tư ít hơn vào phát triển ứng dụng của bạn có nghĩa là sẽ có ROI cao hơn và dự án sẽ có nhiều lợi nhuận hơn

### *Mã nguồn mở*

Lới thế của Android là mã nguồn mở được cấp phép, miễn phí bản quyền và công nghệ tốt nhất được cung cấp bởi cộng đồng Android. Kiến trúc của SDK Android là mã nguồn mở, dễ dàng tương tác với cộng đồng để mở rộng phát triển ứng dụng di động Android. Đây là điều làm cho nền tảng Android trở nên rất hấp dẫn và là cơ hội tốt hơn xây dựng ứng dụng.

### *Triển khai nhanh hơn*

Các ứng dụng Android được viết theo kịch bản bằng ngôn ngữ Java với sự trợ giúp của một bộ thư viện phong phú. Bất cứ ai cũng có thể xây dựng các ứng dụng Android với kiến thức về Java. Việc dễ dàng thích nghi với ngôn ngữ lập trình đó để phát triển ứng dụng di động giúp thích ứng của nó trở nên nhanh hơn và dễ dàng hơn. Android cũng hỗ trợ các ngôn ngữ và công cụ lập trình khác, giúp phát triển ứng dụng có thêm tùy chọn và cách để xây dựng dự án.

### *Tùy chỉnh*

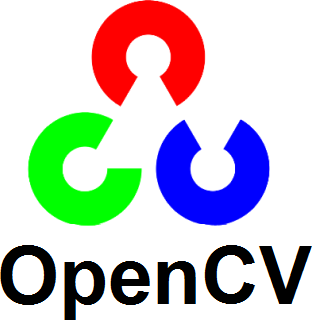
Android là một nền tảng nguồn mở và cung cấp các tính năng tùy chỉnh tối đa cho các nhóm phát triển.Hệ điều hành cho phép tạo ra các ứng dụng Android đa năng có thể dễ dàng tích hợp vào các công cụ đa phương tiện và chức năng quản lý dữ liệu của các quy trình kinh doanh. Nhờ đó dễ dàng sửa chữa, cải thiện và cập nhật ứng dụng. Thêm vào đó có cơ hội điều chỉnh các ứng dụng của mình nhờ lời khuyên của người dùng và đánh giá trong Google Play Store.

### *Tính linh hoạt và khả năng mở rộng*

Với sự xuất hiện của Android Studio, hệ điều hành đã mở rộng về tính linh hoạt và khả năng thích ứng. Nó tích hợp với toàn bộ hệ sinh thái Android bao gồm điện thoại thông minh, máy tính bảng, thiết bị đeo và TV Android. Nó làm cho Ứng dụng Android tương thích với các công nghệ mới nổi như IoT, AR và VR. Do đó, nó cũng là một trong những lợi ích đáng kể của ứng dụng Android. Hơn nữa, tính linh hoạt của nền tảng ứng

dụng Android cho phép các nhóm phát triển xây dựng các ứng dụng di động động phục vụ nhiều mục đích sau khi được cài đặt trên thiết bị.

## OpenCV



*Hình 3.29 - Thư viện OpenCV*

OpenCV là một thư viện mã nguồn mở hàng đầu cho thị giác máy tính (computer vision), xử lý ảnh và máy học, và các tính năng tăng tốc GPU trong hoạt động thời gian thực.

OpenCV được phát hành theo giấy phép BSD, do đó nó hoàn toàn miễn phí cho cả học thuật và thương mại. Nó có các interface C++, C, Python, Java và hỗ trợ Windows, Linux, Mac OS, iOS và Android. OpenCV được thiết kế để tính toán hiệu quả và với sự tập trung nhiều vào các ứng dụng thời gian thực. Được viết bằng tối ưu hóa C/C++, thư viện có thể tận dụng lợi thế của xử lý đa lõi. Được sử dụng trên khắp thế giới, OpenCV có cộng đồng hơn 47 nghìn người dùng và số lượng download vượt quá 6 triệu lần. Phạm vi sử dụng từ nghệ thuật tương tác, cho đến lĩnh vực khai thác mỏ, bản đồ trên web hoặc công nghệ robot.

Thư viện OpenCV bao gồm một số tính năng nổi bật như:

* Bộ công cụ hỗ trợ 2D và 3D
* Nhận diện khuôn mặt
* Nhận diện cử chỉ
* Nhận dạng chuyển động, đối tượng, hành vi,
* Tương tác giữa con người và máy tính
* Điều khiển Robot
* Hỗ trợ thực tế tăng cường

## Kết hợp các công nghệ để xây dựng ứng dụng

* Java là ngôn ngữ lập trình để xây dựng ứng dụng Android.
* Tesseract dùng để xử lý nhận dạng chữ.
* OpenCV để tiền xử lý ảnh.
* AR Core biểu diễn mô hình AR.



*Hình 3.30 - Các công nghệ trong ứng dụng*

## CHƯƠNG 4. PHÂN TÍCH VÀ THIẾT KẾ HỆ THỐNG

## Phân tích yêu cầu

## Sơ đồ use case



*Hình 4.1 - Mô hình use case ứng dụng*

## Danh sách actor

|  |  |  |
| --- | --- | --- |
| **STT** | **Tên Actor** | **Ý nghĩa** |
| **1** | Người dùng | Người dùng sử dụng ứng dụng |

*Bảng 4.1 - Danh sách actor*

## Danh sách use case

|  |  |  |
| --- | --- | --- |
| **STT** | **Tên usecase** | **Mô tả usecase** |
| **1** | Xem giới thiệu tính năng | Cho phép người dùng xem các tính năng nổi bật của ứng dụng |
| **2** | Nhận dạng ký tự | Cho phép người dùng quét ký tự trong vùng nhận dạng được tích hợp trên ứng dụng |
| **3** | Xem danh sách động vật | Cho phép người dùng xem danh sách động vật được tích hợp trên ứng dụng |
| **4** | Trả lời câu hỏi | Cho phép người dùng trả lời các câu hỏi về động vật |
| **5** | Dò tìm mặt phẳng | Cho phép người dùng dò tìm mặt phẳng để dặt mô hình sau khi nhận dạng con vật |
| **6** | Đặt mô hình lên mặt phẳng | Cho phép người dùng đặt mô hình con vật lên mặt phẳng đã tìm |
| **7** | Tương tác mô hình | Cho phép người dùng di chuyển, xoay, phóng to, thu nhỏ mô hình |

|  |  |  |
| --- | --- | --- |
| **8** | Xem hướng dẫn đặt mô hình | Cho phép người dùng xem hướng dẫn cách để đặt mô hình lên mặt phẳng |
| **9** | Thiết lập hiện thông tin nhận dạng ký tự | Cho phép người dùng bật/tắt các thông tin ngôn ngữ, kết quả chữ nhận dạng |
| **10** | Thiết lập thay đổi ngôn ngữ nhận dạng | Cho phép người dùng thay đổi ngôn ngữ nhận dạng |
| **11** | Thiết lập chế độ auto focus | Cho phép người dùng bật/tắt chế độ auto focus của camera |
| **12** | Thiết lập chế độ focus tiêu chuẩn | Cho phép người dùng bật/tắt chế độ focus tiêu chuẩn của camera |
| **13** | Thiết lập tiếng thông báo | Cho phép người dùng bật/tắt tiếng thông báo |
| **14** | Thiết lập âm thanh | Cho phép người dùng bật/tắt âm thanh |

*Bảng 4.2 - Dang sách use case*

## Mô tả chi tiết các use case

* + - 1. **Đặc tả use case “Xem giới thiệu tính năng”**

|  |
| --- |
| **Dòng sự kiện chính** |
| **Mô tả:** Thực hiện xem tính năng nổi bật ứng dụng |
| **Actor:** Người dùng |
| **Điều kiện trước:** Người dùng bắt đầu sử dụng ứng dụng, chưa hoàn tất xem tính năng nổi bật |

|  |  |  |
| --- | --- | --- |
| **Điều kiện sau:** Hoàn tất xem tính năng nổi bật | | |
| **Dòng sự kiện** | | |
| **Bước** | **Hành động người dùng** | **Hành động hệ thống** |
| **1.** | Khởi động ứng dụng |  |
| **2.** |  | Hiện màn hình khởi động  Đi đến màn hình tính năng nổi bật |
| **3.** | Nhấn nút “Next” |  |
| **4.** |  | Chuyển sang màn hình tính năng tiếp theo |
| **5.** | Nhấn nút “Skip” hoặc “Got it” |  |
| **6.** |  | Hoàn tất xem tính năng Đi đến màn hình Home |

*Bảng 4.3- Dòng sự kiện chính use case “Xem giới thiệu tính năng”*

## Đặc tả use case “Nhận dạng ký tự”

|  |
| --- |
| **Dòng sự kiện chính** |
| **Mô tả:** Thực hiện quét ký tự trong vùng nhận dạng |
| **Actor:** Người dùng |
| **Điều kiện trước:** Người dùng đang ở màn hình Home |
| **Điều kiện sau:** Xem được ký tự sau khi nhận dạng |

|  |  |  |
| --- | --- | --- |
| **Dòng sự kiện** | | |
| **Bước** | **Hành động người dùng** | **Hành động hệ thống** |
| **1.** | Ở màn hình Home, nhấn chọn nút “Start” |  |
| **2.** |  | Cấu hình camera  Tạo vùng sáng nhận dạng |
| **3.** | Đưa chữ cần nhận dạng vào vùng sáng. |  |
| **4.** |  | Hiển thị kết quả chữ nhận dạng |

*Bảng 4.4 - Dòng sự kiện chính use case “Nhận dạng ký tự”*

|  |  |  |
| --- | --- | --- |
| **Dòng sự kiện phụ** | | |
| **Mô tả:** Không thể tạo vùng sáng để nhận dạng do cấu hình camera lỗi | | |
| **Actor:** Người dùng | | |
| **Điều kiện trước:** Người dùng đang ở màn hình Home, chưa cấp quyền camera hoặc camera bị hư | | |
| **Điều kiện sau:** Hệ thống hiện thông báo lỗi | | |
| **Dòng sự kiện** | | |
| **Bước** | **Hành động người dùng** | **Hành động hệ thống** |

|  |  |  |
| --- | --- | --- |
| **1.** | Ở màn hình Home, nhấn chọn nút “Start” |  |
| **2.** |  | Hệ thống hiện thông báo lỗi |

*Bảng 4.5 - Dòng sự kiện phụ use case “Nhận dạng ký tự”*

## Đặc tả use case “Xem danh sách động vật”

|  |  |  |
| --- | --- | --- |
| **Dòng sự kiện chính** | | |
| **Mô tả:** Thực hiện xem danh sách các động vật được tích hợp | | |
| **Actor:** Người dùng | | |
| **Điều kiện trước:** Người dùng đang ở màn hình Nhận dạng ký tự | | |
| **Điều kiện sau:** Xem được danh sách các động vật | | |
| **Dòng sự kiện** | | |
| **Bước** | **Hành động người dùng** | **Hành động hệ thống** |
| **1.** | Ở màn hình “Nhận dạng ký tự”, nhấn chọn nút thông tin(?) |  |
| **2.** |  | Load, hiển thị danh sách các động vật |

*Bảng 4.6 - Dòng sự kiện chính use case “Xem danh sách các đông vật”*

## Đặc tả use case “Trả lời câu hỏi”

|  |
| --- |
| **Dòng sự kiện chính** |

|  |  |  |
| --- | --- | --- |
| **Mô tả:** Thực hiện trả lời câu hỏi về động vật | | |
| **Actor:** Người dùng | | |
| **Điều kiện trước:** Người dùng đang ở màn hình Danh sách động vật | | |
| **Điều kiện sau:** Đi đến màn hình Dò mặt phẳng | | |
| **Dòng sự kiện** | | |
| **Bước** | **Hành động người dùng** | **Hành động hệ thống** |
| **1.** | Ở màn hình “Danh sách động vật”, nhấn chọn một động vật |  |
| **2.** |  | Hiện màn hình Câu hỏi |
| **3.** | Nhập câu trả lời đúng và nhấn “Check” |  |
| **4.** |  | Đi đến màn hình Dò mặt phẳng |

*Bảng 4.7- Dòng sự kiện chính use case "Trả lời câu hỏi"*

|  |
| --- |
| **Dòng sự kiện phụ** |
| **Mô tả:** Không nhập câu trả lời hoặc câu trả lời sai |
| **Actor:** Người dùng |
| **Điều kiện trước:** Người dùng đang ở màn hình Câu hỏi |
| **Điều kiện sau:** Hệ thống hiện nhắc nhở người dùng |

|  |  |  |
| --- | --- | --- |
| **Dòng sự kiện** | | |
| **Bước** | **Hành động người dùng** | **Hành động hệ thống** |
| **1.** | Ở vào màn hình “Câu hỏi”, không nhập câu trả lời hoặc nhập câu trả lời sai  Nhấn nút “Check” |  |
| **2.** |  | Hệ thống hiện nhắc nhở người dùng |

*Bảng 4.8 - Dòng sự kiện phụ use case “Trả lời câu hỏi”*

## Đặc tả use case “Dò tìm mặt phẳng”

|  |  |  |
| --- | --- | --- |
| **Dòng sự kiện chính** | | |
| **Mô tả:** Thực hiện dò tìm mặt phẳng để dặt mô hình sau khi nhận dạng con vật | | |
| **Actor:** Người dùng | | |
| **Điều kiện trước:** Người dùng đang ở màn hình Dò mặt phẳng | | |
| **Điều kiện sau:** Xem được mặt phẳng | | |
| **Dòng sự kiện** | | |
| **Bước** | **Hành động người dùng** | **Hành động hệ thống** |
| **1.** | Ở màn hình “Dò mặt phẳng”, di chuyển thiết bị tìm mặt phẳng |  |

|  |  |  |
| --- | --- | --- |
| **2.** |  | Ẩn chỉ dẫn dò tìm  Hiển thị mặt phẳng tìm thấy |

*Bảng 4.9 - Dòng sự kiện chính use case “Dò tìm mặt phẳng”*

|  |  |  |
| --- | --- | --- |
| **Dòng sự kiện phụ** | | |
| **Mô tả:** Không thể tạo được mặt phẳng do điện thoại không hỗ trợ AR | | |
| **Actor:** Người dùng | | |
| **Điều kiện trước:** Người dùng đang ở màn hình Dò mặt phẳng | | |
| **Điều kiện sau:** Hệ thống hiện thông báo lỗi | | |
| **Dòng sự kiện** | | |
| **Bước** | **Hành động người dùng** | **Hành động hệ thống** |
| **1.** | Ở vào màn hình “Dò mặt phẳng”, di chuyển thiết bị tìm mặt phẳng |  |
| **2.** |  | Hệ thống hiện thông báo lỗi |

*Bảng 4.10 - Dòng sự kiện phụ use case “Dò tìm mặt phẳng”*

## Đặc tả use case “Đặt mô hình lên mặt phẳng”

|  |
| --- |
| **Dòng sự kiện chính** |
| **Mô tả:** Thực hiện đặt mô hình con vật lên mặt phẳng đã tìm |
| **Actor:** Người dùng |

|  |  |  |
| --- | --- | --- |
| **Điều kiện trước:** Người dùng đã tìm được mặt phẳng, đang ở màn hình Tương tác mô hình | | |
| **Điều kiện sau:** Mô hình con vật hiện ra | | |
| **Dòng sự kiện** | | |
| **Bước** | **Hành động người dùng** | **Hành động hệ thống** |
| **1.** | Ở màn hình “Tương tác mô hình”, nhấn vào mặt phẳng đã tìm được |  |
| **2.** |  | Hiển thị mô hình AR |

*Bảng 4.11 - Dòng sự kiện chính use case “Đặt mô hình lên mặt phẳng”*

## Đặc tả use case “Tương tác mô hình”

|  |  |  |
| --- | --- | --- |
| **Dòng sự kiện chính** | | |
| **Mô tả:** Thực hiện di chuyển, xoay, phóng to, thu nhỏ mô hình | | |
| **Actor:** Người dùng | | |
| **Điều kiện trước:** Người dùng đang ở màn hình Tương tác mô hình | | |
| **Điều kiện sau:** Mô hình di chuyển, xoay, phóng to, thu nhỏ theo điều chỉnh | | |
| **Dòng sự kiện** | | |
| **Bước** | **Hành động người dùng** | **Hành động hệ thống** |
| **1.** | Ở màn hình “Tương tác mô hình”, nhấn chọn mô hình |  |

|  |  |  |
| --- | --- | --- |
| **2.** |  | Hiển thị vòng sáng quanh mô hình |
| **3.** | Nhấn giữ di chuyển ngón tay và thả ra ở vị trí mới; Nhấn giữ dùng hai ngón tay kéo thu hẹp hoặc mở rộng; Nhấn giữ dùng hai ngón tay xoay mô hình |  |
| **4.** |  | Mô hình chuyển sang vị trí được chọn; Mô hình thu nhỏ hoặc phóng to theo điều chỉnh; Mô hình xoay theo điều chỉnh |

*Bảng 4.12 - Dòng sự kiện chính use case “Tương tác mô hình”*

## Đặc tả use case “Xem hướng dẫn đặt mô hình”

|  |  |  |
| --- | --- | --- |
| **Dòng sự kiện chính** | | |
| **Mô tả:** Thực hiện xem hướng dẫn cách để xuất hiện mô hình | | |
| **Actor:** Người dùng | | |
| **Điều kiện trước:** Người dùng đang ở màn hình Dò mặt phẳng | | |
| **Điều kiện sau:** Xem được hướng dẫn | | |
| **Dòng sự kiện** | | |
| **Bước** | **Hành động người dùng** | **Hành động hệ thống** |

|  |  |  |
| --- | --- | --- |
| **1.** | Ở vào màn hình “Dò mặt phẳng”, nhấn chọn nút thông tin(?) |  |
| **2.** |  | Hiển thị hướng dẫn |

*Bảng 4.13 - Dòng sự kiện chính use case “Xem hướng dẫn đặt mô hình”*

## Đặc tả use case “Thiết lập hiện thông tin nhận dạng ký tự”

|  |  |  |
| --- | --- | --- |
| **Dòng sự kiện chính** | | |
| **Mô tả:** Thực hiện bật/tắt các thông tin ngôn ngữ, kết quả chữ nhận dạng | | |
| **Actor:** Người dùng | | |
| **Điều kiện trước:** Người dùng đang ở màn hình Cài đặt | | |
| **Điều kiện sau:** Thấy thông tin ngôn ngữ nhận dạng, kết quả chữ nhận dạng | | |
| **Dòng sự kiện** | | |
| **Bước** | **Hành động người dùng** | **Hành động hệ thống** |
| **1.** | Ở màn hình “Cài đặt”, nhấn chọn vào ô mục “Show result” |  |
| **2.** |  | Hệ thống ghi nhận việc bật/tắt thông tin nhận dạng ký tự |
| **3.** | Nhấn mũi bên (back) trở về màn hình Home  Nhấn chọn nút “Start” |  |

|  |  |  |
| --- | --- | --- |
| **4.** |  | Hệ thống cho phép thấy thông tin nhận dạng chữ |

*Bảng 4.14 - Dòng sự kiện chính use case “Thiết lập hiện thông tin nhận dạng chữ”*

## Đặc tả use case “Thiết lập thay đổi ngôn ngữ nhận dạng”

|  |  |  |
| --- | --- | --- |
| **Dòng sự kiện chính** | | |
| **Mô tả:** Thực hiện thay đổi ngôn ngữ nhận dạng | | |
| **Actor:** Người dùng | | |
| **Điều kiện trước:** Người dùng đang ở màn hình Cài đặt | | |
| **Điều kiện sau:** Thay đổi được ngôn ngữ | | |
| **Dòng sự kiện** | | |
| **Bước** | **Hành động người dùng** | **Hành động hệ thống** |
| **1.** | Ở màn hình “Cài đặt”, nhấn chọn vào mục “Recognize” |  |
| **2.** |  | Hệ thị danh sách các ngôn ngữ |
| **3.** | Chọn ngôn ngữ |  |
| **4** |  | Hệ thống thiết lập chuyển sang ngôn ngữ được chọn |

|  |  |  |
| --- | --- | --- |
| **5.** | Nhấn mũi bên (back) trở về màn hình Home  Nhấn chọn nút “Start” |  |
| **6.** |  | Cập nhật thay đổi ngôn ngữ |

*Bảng 4.15 - Dòng sự kiện chính use case “Thiết lập thay đổi ngôn ngữ nhận dạng”*

|  |  |  |
| --- | --- | --- |
| **Dòng sự kiện phụ** | | |
| **Mô tả:** Không thể tải ngôn ngữ do không có kết nối mạng hoặc tải thất bại | | |
| **Actor:** Người dùng | | |
| **Điều kiện trước:** Người dùng đang ở màn hình Home, không có kết nối mạng | | |
| **Điều kiện sau:** Hệ thống hiện thông báo lỗi | | |
| **Dòng sự kiện** | | |
| **Bước** | **Hành động người dùng** | **Hành động hệ thống** |
| **1.** | Ở màn hình Home, nhấn chọn nút “Start” |  |
| **2.** |  | Hệ thống hiện thông báo lỗi |

*Bảng 4.16 - Dòng sự kiện phụ use case “Thiết lập thay đổi ngôn ngữ nhận dạng”*

## Đặc tả use case “Thiết lập chế độ auto focus”

|  |  |  |
| --- | --- | --- |
| **Dòng sự kiện chính** | | |
| **Mô tả:** Thực hiện bật/tắt chế độ auto focus của camera | | |
| **Actor:** Người dùng | | |
| **Điều kiện trước:** Người dùng đang ở màn hình Cài đặt | | |
| **Điều kiện sau:** Thay đổi thành công | | |
| **Dòng sự kiện** | | |
| **Bước** | **Hành động người dùng** | **Hành động hệ thống** |
| **1.** | Ở màn hình “Cài đặt”, nhấn chọn vào ô mục “Autofocus” |  |
| **2.** |  | Hệ thống thiết lập chế độ camera tự động lấy nét |

*Bảng 4.17 - Dòng sự kiện chính use case “Thiết lập chế độ auto focus”*

## Đặc tả use case “Thiết lập chế độ focus tiêu chuẩn”

|  |
| --- |
| **Dòng sự kiện chính** |
| **Mô tả:** Thực hiện bật/tắt chế độ focus tiêu chuẩn của camera |
| **Actor:** Người dùng |
| **Điều kiện trước:** Người dùng đang ở màn hình Cài đặt |
| **Điều kiện sau:** Thay đổi thành công |

|  |  |  |
| --- | --- | --- |
| **Dòng sự kiện** | | |
| **Bước** | **Hành động người dùng** | **Hành động hệ thống** |
| **1.** | Ở màn hình “Cài đặt”, nhấn chọn vào ô mục “Standard focus mode” |  |
| **2.** |  | Hệ thống thiết lập chế độ camera lấy nét tiêu chuẩn |

*Bảng 4.18 - Dòng sự kiện chính use case “Thiết lập chế độ focus tiêu chuẩn”*

## Đặc tả use case “Thiết lập tiếng thông báo”

|  |  |  |
| --- | --- | --- |
| **Dòng sự kiện chính** | | |
| **Mô tả:** Thực hiện bật/tắt tiếng thông báo | | |
| **Actor:** Người dùng | | |
| **Điều kiện trước:** Người dùng đang ở màn hình Cài đặt | | |
| **Điều kiện sau:** Thay đổi thành công | | |
| **Dòng sự kiện** | | |
| **Bước** | **Hành động người dùng** | **Hành động hệ thống** |
| **1.** | Ở màn hình “Cài đặt”, nhấn chọn vào ô mục “Beep” |  |
| **2.** |  | Hệ thống thiết lập bật/tắt tiếng thông báo |

*Bảng 4.19 - Dòng sự kiện chính use case “Thiết lập tiếng thông báo”*

## Đặc tả use case “Thiết lập âm thanh”

|  |  |  |
| --- | --- | --- |
| **Dòng sự kiện chính** | | |
| **Mô tả:** Thực hiện bật/tắt âm thanh | | |
| **Actor:** Người dùng | | |
| **Điều kiện trước:** Người dùng đang ở màn hình Cài đặt | | |
| **Điều kiện sau:** Thay đổi thành công | | |
| **Dòng sự kiện** | | |
| **Bước** | **Hành động người dùng** | **Hành động hệ thống** |
| **3.** | Ở màn hình “Cài đặt”, nhấn chọn vào ô mục “Music” |  |
| **4.** |  | Hệ thống thiết lập bật/tắt âm thanh |

*Bảng 4.20 - Dòng sự kiện chính use case “Thiết lập âm thanh”*

## Thiết kế dữ liệu

## Sơ đồ các thành phần chính



*Hình 4.2 - Sơ đồ các thành phần chính*

## Mô tả các thành phần có trong ứng dụng

Danh sách các các thành phần chính.

|  |  |  |
| --- | --- | --- |
| STT | Tên module | Mô tả |
| 1 | Camera Manager | Quản lý cấu hình, hoạt động camera |
| 2 | Handler ORC | Xử lý nhận dạng chữ |
| 3 | Sound Manager | Quản lý âm thanh |
| 4 | AR Manager | Xử lý thực tế ảo tăng cường AR |
| 5 | Cogfig app | Quản lý cấu hình ứng dụng |

*Bảng 4.21 - Các thành phần chính trong cơ sở ứng dụng*

## Camera Manager

* + - Camera Config: cấu hình các parameters của camera (kích thước ống ngắm, đèn, focus,...). Dựa vào screen resolution (phân giải màn hình) tìm thông số phù hợp cho camera resolution (độ phân giải camera).
    - Preview Callback : Nhận một xử lý preview và gửi thông điệp phản hồi (message) bao gồm camera resolution và data (hình ảnh thu được) cho camera.
    - Auto Focus: Tùy chỉnh lấy nét tự động bao gồm chế độ lấy nét (focus mode), thời gian lấy nét (focus interval).
    - Thực hiển đóng / mở camera.
    - Thực hiện bắt đầu / kết thúc preview.
    - Yêu cầu giải mã ORC.
    - Yêu cầu chạm lấy nét.
    - Tạo vùng sáng quét chữ (luminan source).

## Handler ORC

* + - Decode Thread: luồng giải mã với các xử lý (decode handler). Mỗi xử lý nhận hình ảnh (chiều rộng, chiều cao) và API Tesseraact thực hiện xử lý bất đồng bộ (AsyncTask). Sử dụng vùng sáng để lấy thông tin bitmap, kết hợp API Tesseraact, sau đó giải mã nhận được kết quả (ORC result) bao gồm text và thời gian thực hiện.
    - Quản lý trạng thái (state): Preview, Preview paused, continous, continous paused, Succcess, Done.
    - Xử lý các hành vi: decode, decode succeeded, decode failed, continuous decode, continuous decode succeeded, continuous decode failed, ….

## Sound Manager

* + - Âm lượng tiếng thông báo, nhạc nền.
    - Quản lý trạng thái: bắt đầu, tạm dừng, tắt.

## AR Manager

* + - Xử lý phát hiện bề mặt.
    - Dựng mô hình AR.
    - Tạo điểm neo cho mô hình trong tọa độ không gian 3D.
    - Tạo tương tác lên mô hình: xoay, phóng to, thu nhỏ.

## Config App

* + - Điều chỉnh hiện thị thông tin nhận dạng:ngôn ngữ, kết quả.
    - Thay đổi ngôn ngữ nhận dạng

## Danh sách các ngôn ngữ nhận dạng

|  |  |  |
| --- | --- | --- |
| **STT** | **Tên ngôn ngữ** | **Viết tắt chuẩn iso6393** |
| 1 | Afrikaans | afr |
| 2 | Albanian | sqi |
| 3 | Arabic | ara |
| 4 | Azeri | aze |
| 5 | Basque | eus |
| 6 | Belarusian | bel |
| 7 | Bengali | ben |
| 8 | Bulgarian | bul |
| 9 | Catalan | cat |
| 10 | Chinese (Simplified) | chi\_sim |

|  |  |  |
| --- | --- | --- |
| 11 | Chinese (Traditional) | chi\_tra |
| 12 | Croatian | hrv |
| 13 | Czech | ces |
| 14 | Danish | dan |
| 15 | Dutch | nld |
| 16 | English | eng |
| 17 | Estonian | est |
| 18 | Finnish | fin |
| 19 | French | fra |
| 20 | Galician | glg |
| 21 | German | deu |
| 22 | Greek | ell |
| 23 | Hebrew | heb |
| 24 | Hindi | hin |
| 25 | Hungarian | hun |
| 26 | Icelandic | isl |
| 27 | Indonesian | ind |
| 28 | Italian | ita |
| 29 | Japanese | jpn |

|  |  |  |
| --- | --- | --- |
| 30 | Kannada | kan |
| 31 | Korean | kor |
| 32 | Latvian | lav |
| 33 | Lithuanian | lit |
| 34 | Macedonian | mkd |
| 35 | Malay | msa |
| 36 | Malayalam | mal |
| 37 | Maltese | mlt |
| 38 | Norwegian | nor |
| 39 | Polish | pol |
| 40 | Portuguese | por |
| 41 | Romanian | ron |
| 42 | Russian | rus |
| 43 | Serbian | srp |
| 44 | Slovak | slk |
| 45 | Slovenian | slv |
| 46 | Spanish | spa |
| 47 | Swahili | swa |
| 48 | Swedish | swe |

|  |  |  |
| --- | --- | --- |
| 49 | Tagalog | tgl |
| 50 | Tamil | tam |
| 51 | Telugu | tel |
| 52 | Thai | tha |
| 53 | Turkish | tur |
| 54 | Ukrainian | ukr |
| 55 | Vietnamese | vie |

*Bảng 4.22 - Danh sách các ngôn ngữ nhận dạng*

## Danh sách các động vật

|  |  |  |
| --- | --- | --- |
| **STT** | **Tên tiếng Anh** | **Tên tiếng Việt** |
| 1 | armadillo | con tatu |
| 2 | bear | gấu |
| 3 | beaver | hải ly |
| 4 | bee | ong |
| 5 | bird | chim |
| 6 | bison | bò rừng |
| 7 | butterfly | bướm |
| 8 | camel | lạc đà |

|  |  |  |
| --- | --- | --- |
| 9 | cat | mèo |
| 10 | chicken | gà |
| 11 | cow | bò |
| 12 | crab | cua |
| 13 | crocodile | cá sấu |
| 14 | deer | nai |
| 15 | dinosaur | khủng long |
| 16 | dog | chó |
| 17 | dolphin | cá heo |
| 18 | duck | vịt |
| 19 | elephant | voi |
| 20 | ferret | chồn |
| 21 | fish | cá |
| 22 | fox | cáo |
| 23 | frog | ếch |
| 24 | gibbon | vượn |
| 25 | giraffe | hươu cao cổ |
| 26 | goat | dê |
| 27 | goose | ngỗng |

|  |  |  |
| --- | --- | --- |
| 28 | gull | mòng biển |
| 29 | hawk | chim ưng |
| 30 | hippopotamus | hà mã |
| 31 | horse | ngựa |
| 32 | hyena | linh cẩu |
| 33 | kangaroo | chuột túi |
| 34 | kingfisher | chim bói cá |
| 35 | koala | gấu koala |
| 36 | lamb | cừu |
| 37 | lion | sưu tử |
| 38 | lizard | thằn lằn |
| 39 | mammoth | voi ma mút |
| 40 | manatee | lợn biển |
| 41 | monkey | khỉ |
| 42 | otter | rái cá |
| 43 | panda | gấu trúc |
| 44 | parrot | con vẹt |
| 45 | peacock | con công |
| 46 | penguin | chim cánh cụt |

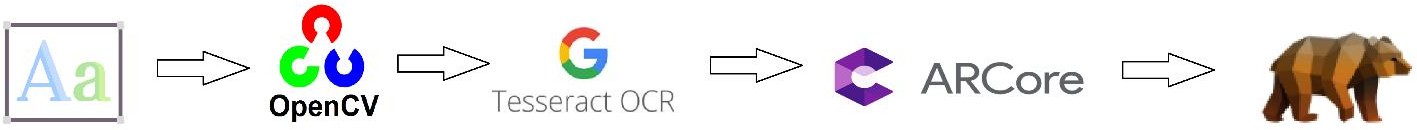
|  |  |  |
| --- | --- | --- |
| 47 | pig | lợn |
| 48 | rabbit | thỏ |
| 49 | racoon | gấu mèo |
| 50 | reindeer | tuần lộc |
| 51 | seahorse | cá ngựa |
| 52 | sea lion | sư tử biển |
| 53 | shark | cá mập |
| 54 | sheep | cừu |
| 55 | shrimp | tôm |
| 56 | snail | ốc sên |
| 57 | snake | con rắn |
| 58 | squirrel | con sóc |
| 59 | stork | con cò |
| 60 | swan | thiên nga |
| 61 | tapir | heo vòi |
| 62 | tiger | hổ |
| 63 | turtle | rùa |
| 64 | mouse | chuột |
| 65 | vulture | kền kền |

|  |  |  |
| --- | --- | --- |
| 66 | walrus | hải mã |
| 67 | whale | cá voi |
| 68 | wolf | chó sói |
| 69 | wolverine | chồn sói |

*Bảng 4.23 - Danh sách các động vật*

## Thiết kế kiến trúc

## Sơ đồ kiến trúc tổng thể

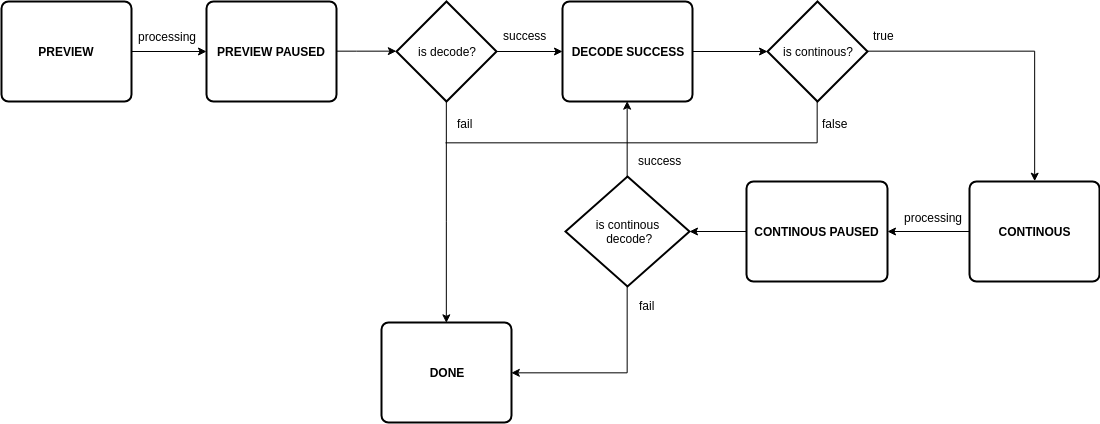


*Hình 4.3 - Mô hình kiến trúc tổng thể*

Hệ thống bao gồm 3 thành phần chính và tương giác giữa chúng như sau:

* Bộ xử lý ảnh sử dụng openCV để chuyển màu ảnh sang trắng đen giúp cho việc nhận dạng ký tự hiểu quả hơn.
* Bộ xử lý nhận dạng ký tự với Tesseract OCR sử sụng LSTM.
* Bộ xử lý mô hình sử dụng ARCore tạo mô hình động vật dạng thực tế ảo.

## Sơ đồ quản lý trạng thái

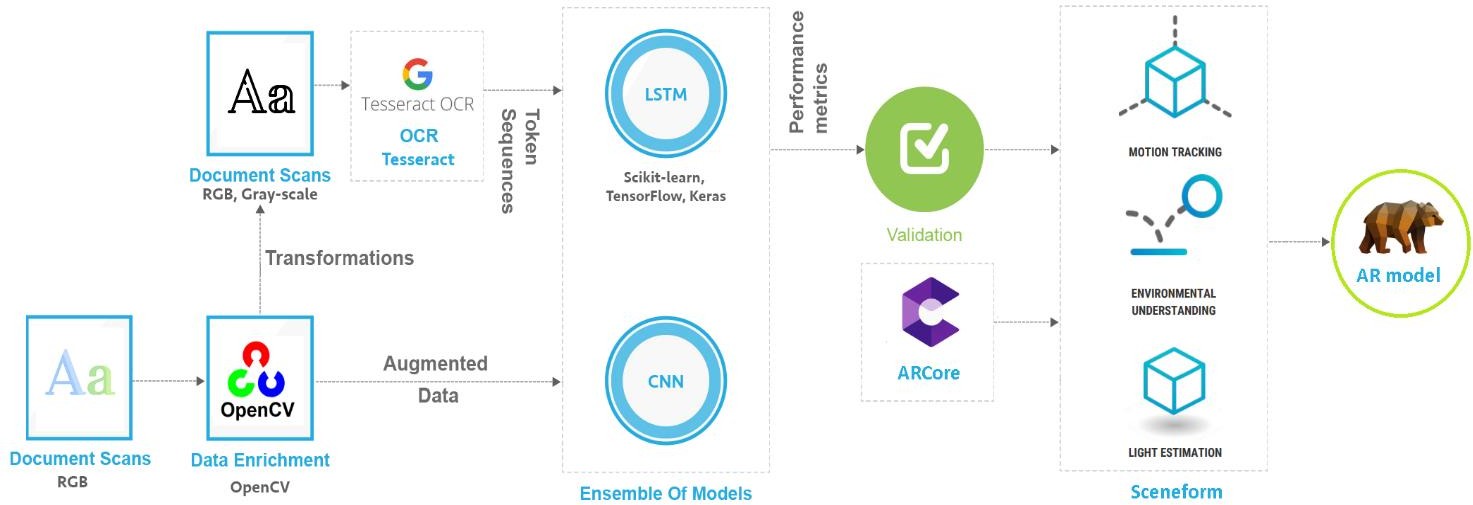


*Hình 4.4 - Sơ đồ hoạt động nhận dạng*

Các trang thái hoạt động trong sơ đồ:

* PREVIEW: bắt đầu quá trình nhận dạng, xử lý tìm ký tự.
* PREVIEW PAUSE: tìm thấy ký tự, xử lý nhận dạng.
* DECODE SUCCESS: nhận dạng ký tự thành công
* CONTINOUS: xử lý tiếp tục quá trình nhận dạng
* CONTINOUS PAUSED: tìm thấy ký tự, xử lý nhận dạng
* DONE: kết thúc quá trình nhận dạng Các lưồng trạng thái miêu tả như sau:
* Người dùng bắt đầu quá trình nhận dạng, ứng dụng xử lý tìm ký tự.
* Mỗi khi tìm thấy ký tự, xử lý nhận dạng ký tự theo ngôn ngữ sẽ được thưc hiện đa luồng.
* Sau khi xử lý, nếu kết quả thành công, trả về ký tự nhận dạng được và chuyển sang quá trình xử lý tiếp tục, nếu thất bại kết thúc luồng của quá trình nhận dạng.

## Sơ đồ kiến trúc chi tiết



*Hình 4.5 - Sơ đồ kiến trúc chi tiết*

Chi tiết kiến trúc:

1. Bộ xử lý ảnh
   * Document Scans: ảnh màu chứa ký tự chưa qua xử lý
   * Document Scans (Gray): ảnh đã qua xử lý thành ảnh trắng đen
   * Thư viện OpenCV: dùng để xử lý ảnh
2. Bộ xử lý nhận dạng ký tự
   * Thư viện Tesseract OCR: dùng nhận dạng ký tự.
   * RNN (Recurrent Neural Network) Mạng nơ-ron hồi quy chứa các vòng lặp bên trong cho phép thông tin có thể lưu lại được.
   * LSTM (Long Short Term Memory networks) một dạng đặc biệt của RNN để dự đoán mô tả từ tất cả các từ đầu vào.
3. Bộ xử lý mô hình
   * Bộ công cụ ARCore tạo mô hình động vật dạng thực tế ảo.

+ Motion tracking giúp điện thoại hiểu được và theo dõi được vị trí của nó so với thế giới thực.

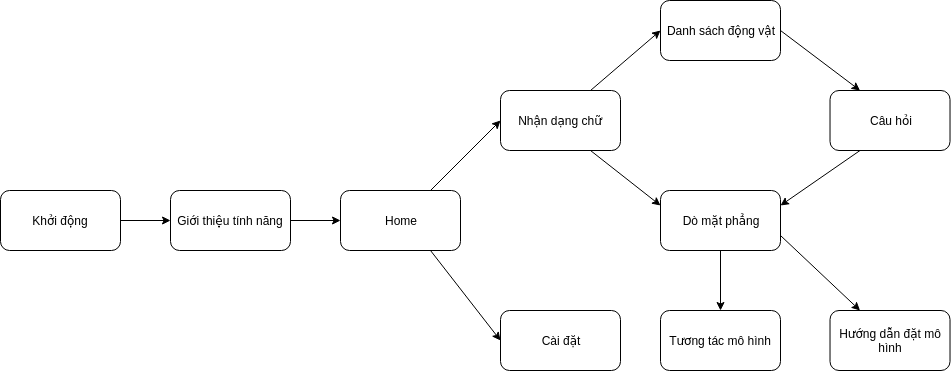
+ Environmental understanding giúp điện thoại phát hiện được kích cỡ và vị trí của tất cả các bề mặt: ngang, dọc và các góc cạnh bề mặt như mặt đất, mặt bàn hay bức tường...

+ Light estimation giúp điện thoại ước lượng được điều kiện ánh sáng của môi trường.

* + AR model: mô hình thực tế ảo tăng cường.

## Thiết kế giao diện

## Sơ đồ thiết kế giao diện



*Hình 4.6 - Sơ đồ thiết kế giao diện*

## Danh sách giao diện

|  |  |  |  |
| --- | --- | --- | --- |
| STT | Tên giao diện | Mô tả giao diện | Ghi chú |
| 1. | Khởi động | Giao diện lúc khởi động ứng dụng |  |
| 2. | Giới thiệu tính năng | Giao diện giới thiệu các tính năng nổi  bật của ứng dụng. |  |
| 3. | Home | Giao diện chính |  |
| 4. | Nhận dạng chữ | Giao diện quét camera nhận dạng chữ |  |

|  |  |  |  |
| --- | --- | --- | --- |
| 5. | Danh sách động vật | Giao diện danh các loài động vật tích  hợp trong ứng dụng |  |
| 6. | Câu hỏi | Giao diện câu hỏi về tên động vật |  |
| 7. | Dò mặt phẳng | Giao diện dò tìm mặt phẳng để đặt con  vật sau khi nhận dạng |  |
| 8. | Tương tác mô hình | Giao diện tương tác các mô hình AR |  |
| 9. | Hướng dẫn đặt mô hình | Giao diện hướng dẫn cách đặt mô hình  AR lên mặt phẳng |  |
| 10. | Cài đặt | Giao diện cài đặt ứng dụng |  |

*Bảng 4.24 - Danh sách các giao diện*

## Mô trả chi tiết giao diện

* + - 1. Giao diện Khởi động

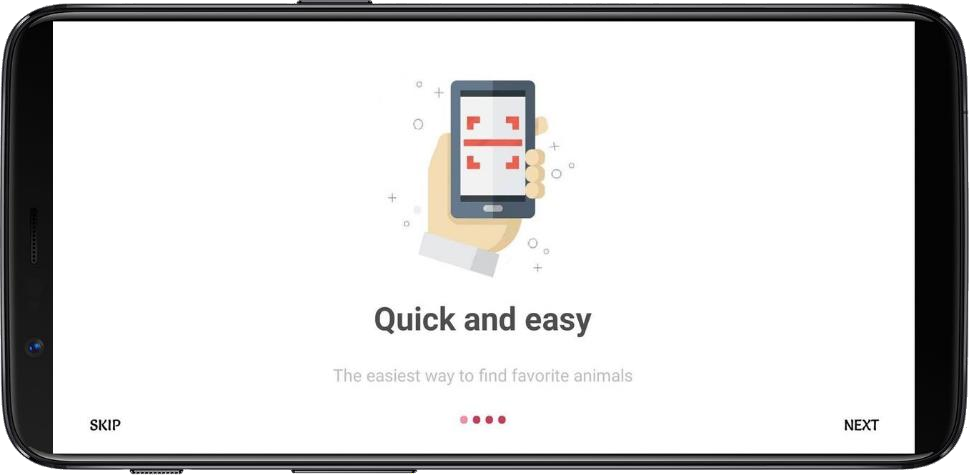


*Hình 4.7 - Giao diện “Khởi động”*

|  |  |  |  |
| --- | --- | --- | --- |
| STT | Tên thành phần | Loại thành phần | Mô tả chức năng của thành phần |
| 1 | Logo | Image | Logo ứng dụng |
| 2 | Domain | Text | Đường dẫn trang web nhà phát hành |

*Bảng 4.25 - Thành phần giao diện “Khởi động”*

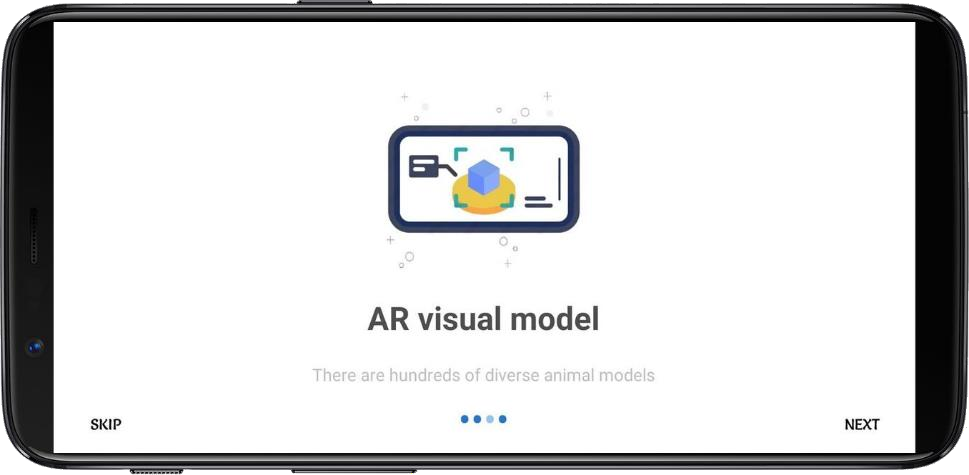
* + - 1. Giao diện Giới thiệu tính năng



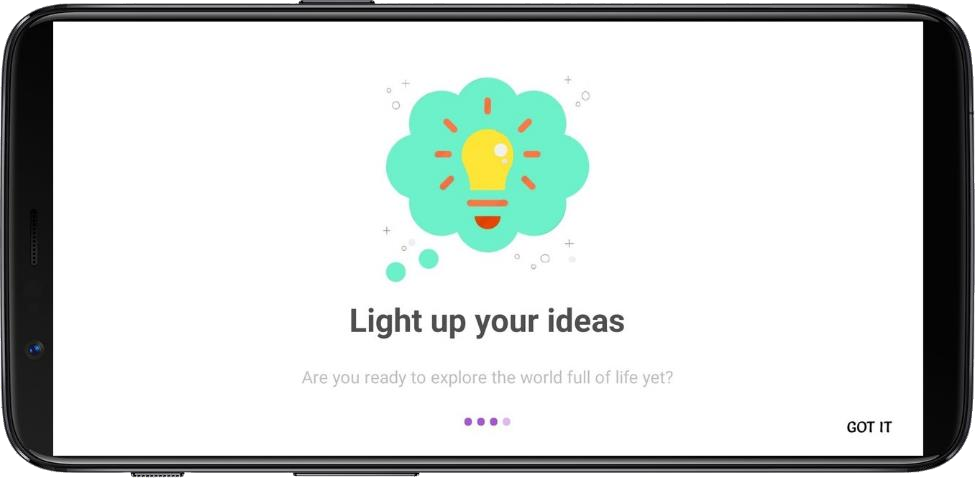
*Hình 4.8 - Giao diện “Giới thiệu tính năng”*



*Hình 4.9 - Giao diện “Giới thiệu tính năng”*



*Hình 4.10 - Giao diện “Giới thiệu tính năng”*



*Hình 4.11 - Giao diện “Giới thiệu tính năng”*

|  |  |  |  |
| --- | --- | --- | --- |
| STT | Tên thành phần | Loại thành phần | Mô tả chức năng của thành phần |
| 1 | SKIP | Button | Bỏ qua xem tính năng, đi đến giao  màn hình Home |
| 2 | NEXT | Button | Đi đến màn hình tính năng tiếp theo |
| 3 | GOT IT | Button | Hoàn tất xem tính năng, đi đến màn  hình Home |

*Bảng 4.26 - Thành phần giao diện “Giới thiệu tính năng”*

* + - 1. Giao diện Home

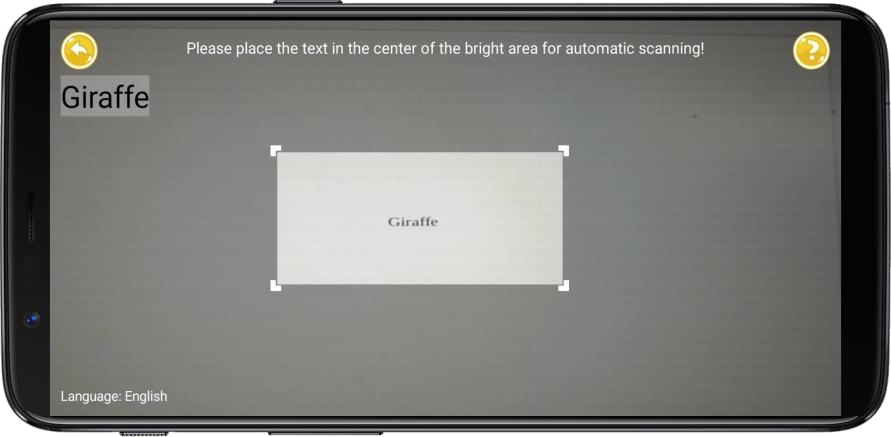


*Hình 4.12 - Giao diện “Home”*

|  |  |  |  |
| --- | --- | --- | --- |
| STT | Tên thành phần | Loại thành phần | Mô tả chức năng của thành phần |
| 1 | Nút START | Button | Đi đến màn hình Nhận dạng chữ |
| 2 | Nút SETTING | Button | Đi đến màn hình Cài đặt |
| 3 | Nút QUIT | Button | Thoát khỏi ứng dụng |

*Bảng 4.27 – Thành phần giao diện “Home”*

* + - 1. Giao diện Nhận dạng chữ

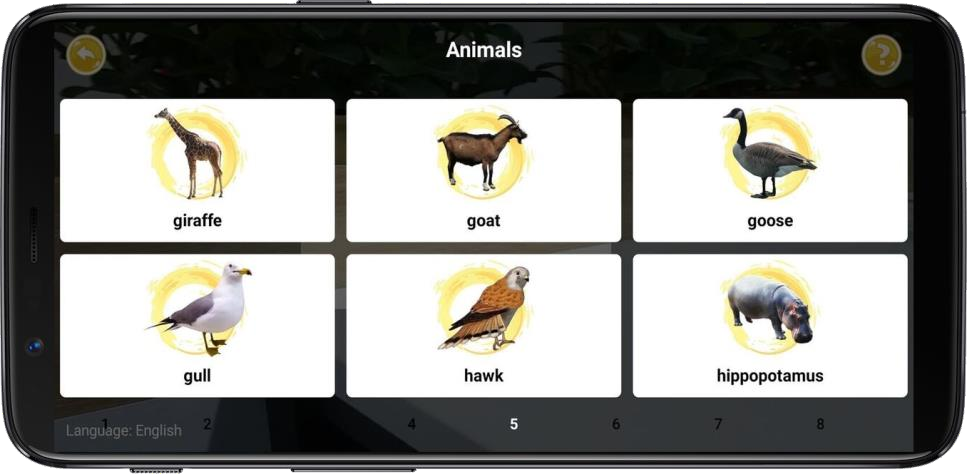


*Hình 4.13 - Giao diện nhận dạng chữ*

|  |  |  |  |
| --- | --- | --- | --- |
| STT | Tên thành phần | Loại thành phần | Mô tả chức năng của thành phần |
| 1 | Nút mũi tên | Button | Trở về màn hình Home |
| 2 | Nút thông tin (?) | Buttosn | Đi đến màn hình Danh sách động vật |
| 3 | Kết quả nhận  dạng | Text | Kế quả nhận dạng ký tự |
| 4 | Ngôn ngữ nhận  dạng (Language) | Text | Ngôn ngữ được sử dụng để nhận dạng  ký tự |
| 5 | Vùng sáng nhận  dạng | Bitmap | Lấy thông tin từ camera chuyển thành  ảnh để nhận dạng |

*Bảng 4.28 – Thành phần giao diện “Nhận dạng chữ”*

* + - 1. Giao diện Danh sách động vật



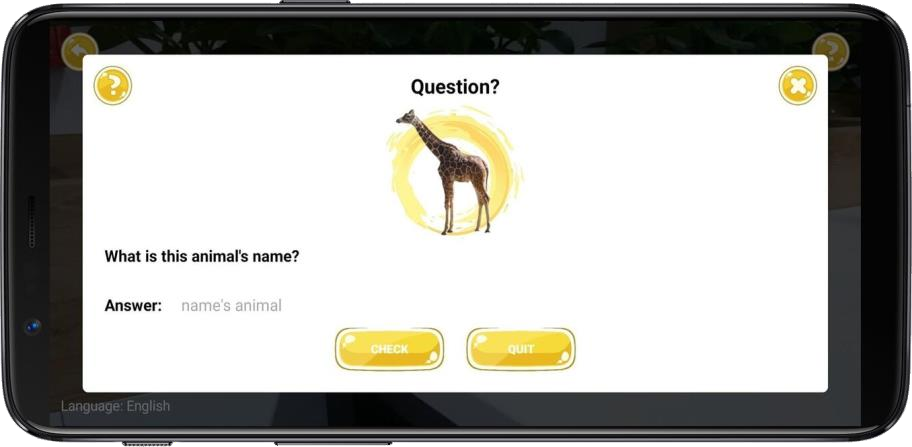
*Hình 4.14 - Giao diện “Danh sách động vật”*

|  |  |  |  |
| --- | --- | --- | --- |
| STT | Tên thành phần | Loại thành phần | Mô tả chức năng của thành phần |
| 1 | Thẻ động vật | Image, Text | Hình ảnh và tên động vật |

|  |  |  |  |
| --- | --- | --- | --- |
| 2 | Phân trang | List button | Chuyển đổi danh sách theo xử lý phân  trang |

*Bảng 4.29 – Thành phần giao diện “Danh sách động vật”*

* + - 1. Giao diện Câu hỏi



*Hình 4.15 - Giao diện “Câu hỏi”*

|  |  |  |  |
| --- | --- | --- | --- |
| STT | Tên thành phần | Loại thành phần | Mô tả chức năng của thành phần |
| 1 | Nút đóng (X) | Button | Đóng màn hình câu hỏi |
| 2 | Animal’s name | Edittext | Nhập tên con vật |
| 3 | CHECK | Button | Kiểm tra tên con vật |
| 4 | QUIT | Button | Đóng màn hình câu hỏi |

*Bảng 4.30 – Thành phần giao diện “Câu hỏi”*

* + - 1. Giao diện Dò mặt phẳng



*Hình 4.16 - Giao diện “Dò mặt phẳng”*

|  |  |  |  |
| --- | --- | --- | --- |
| STT | Tên thành phần | Loại thành phần | Mô tả chức năng của thành phần |
| 1 | Nút mũi tên | Button | Trở về màn hình Nhận dạng chữ |
| 2 | Nút thông tin (?) | Button | Đi đến màn hình Hướng dẫn đặt mô  hình |
| 3 | Hoạt ảnh hướng  dẫn | Animation | Gợi ý cách tìm mặt phẳng |

*Bảng 4.31 – Thành phần giao diện “Dò mặt phẳng”*

* + - 1. Giao diện Tương tác mô hình



*Hình 4.17 - Giao diện “Tương tác mô hình”*

|  |  |  |  |
| --- | --- | --- | --- |
| STT | Tên thành phần | Loại thành phần | Mô tả chức năng của thành phần |
| 1 | Nút mũi tên | Button | Trở về màn hình Nhận dạng chữ |
| 2 | Nút thông tin (?) | Button | Đi đến màn hình Hướng dẫn đặt mô  hình |
| 3 | Mô hình AR | Model | Mô hình động vật cho phép di  chuyển, xoay, phóng to, thu nhỏ |

*Bảng 4.32 - Thành phần giao diện “Tương tác mô hình”*

* + - 1. Giao diện Hướng dẫn đặt mô hình

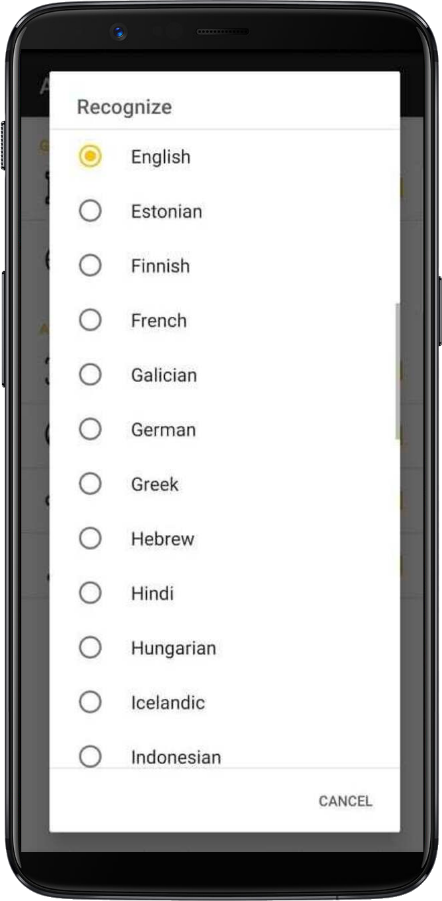
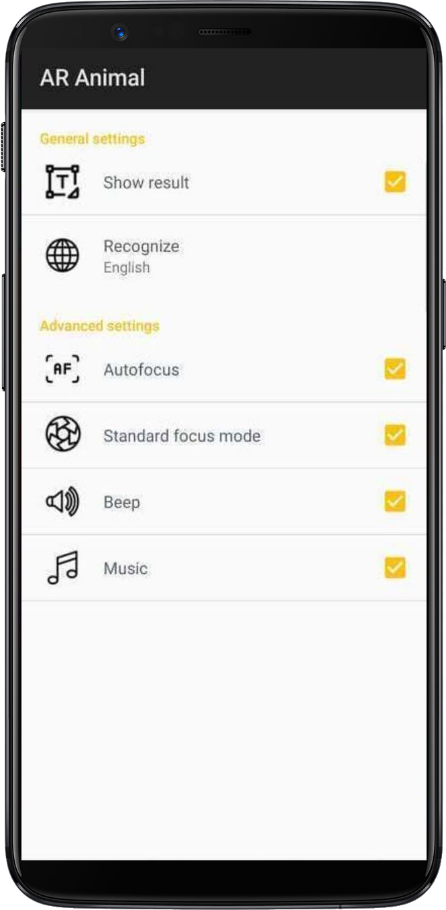


*Hình 4.18 - Giao diện “Hướng dẫn đặt mô hình”*

|  |  |  |  |
| --- | --- | --- | --- |
| STT | Tên thành phần | Loại thành phần | Mô tả chức năng của thành phần |
| 1 | Ảnh hướng dẫn | Image | Hình ảnh hướng dẫn cách đặt mô hình |

*Bảng 4.33 - Thành phần giao diện “Hướng dẫn đặt mô hình”*

* + - 1. Giao diện cài đặt



*Hình 4.19 - Giao diện “Cài đặt”*

|  |  |  |  |
| --- | --- | --- | --- |
| STT | Tên thành phần | Loại thành phần | Mô tả chức năng của thành phần |
| 1 | Show result | Checkbox | Cho phép bật/tắt thông tin kết quả  nhận dạng, ngôn ngữ nhận dạng |
| 2 | Recognize | List | Chọn ngôn ngữ nhận dạng |
| 3 | Autofocus | Checkbox | Cho phép bật/tắt chế độ autofocus  camera |
| 4 | Standard focus  mode | Checkbox | Cho phép bật/tắt chê độ focus tiêu  chuẩn camare |

|  |  |  |  |
| --- | --- | --- | --- |
| 5 | Beep | Checkbox | Cho phép bật/tắt tiếng thông báo |
| 6 | Music | Checkbox | Cho phép bật/tắt âm thanh |

*Bảng 4.34 - Thành phần giao diện “Hướng dẫn đặt mô hình”*

# CHƯƠNG 5. KẾT LUẬN VÀ HƯỚNG PHÁT TRIỂN

## Kết quả đạt được

Trong quá trình thực hiện khóa luận tốt nghiệp, nhóm chúng em đã có thêm điều kiện tìm hiểu công nghệ mới, cách xây dựng một ứng dụng cho lĩnh vực nghiên cứu về dạng mẫu, trí tuệ nhân tạo và machine vision, thực tế ảo tăng cường cùng với đó là kinh nghiệm lên kế hoạch tìm hiểu, nghiên cứu phát triển đề tài, viết báo cáo... Đồng thời, chúng em có thể củng cố và có thêm nhiều kinh nghiệm trong quá trình xây dựng và tổ chức mã nguồn của ứng dụng.

Qua quá trình nghiên cứu và phát triển đề tài, nhóm chúng em đã hoàn thành được các tính năng sau:

* Có đầy đủ những chức năng đã đề ra.
* Hỗ trợ UI thân thiện và bắt mắt cho người dùng.
* Hỗ trợ UX mang lại trải nghiệm tốt nhất cho người dùng.

## Nhận xét

## Ưu điểm

* + - * Dễ sử dung, giao diên thân thiện.
      * Nhận dạng ký tự nhanh, chính xác.
      * Hỗ trợ nhiều động vật và ngôn ngữ.
      * Áp dụng các công nghệ mới vào ứng dụng.

## Khuyết điểm

* + - * Nhận dạng ký tự trong điều kiện thiếu sáng chưa tốt.
      * Quá trình dò mặt phẳng đôi lúc khá chậm.
      * Chưa hỗ trợ âm thanh, chuyển động của con vật.

## Hướng phát triển

Đề tài cho phép chúng ta mở rộng và phát triển nhằm xây dựng một ứng dụng hoàn chỉnh:

* + Xử lý nhận dạng trong điều kiện thiếu sáng.
  + Cải thiện quá trình tìm mặt phẳng.
  + Hỗ trợ tương tác với mô hình.
  + Đa dạng các động vật và ngôn ngữ nhận dạng.
  + Chia sẻ hình ảnh, video quay các mô hình.

# TÀI LIỆU THAM KHẢO

1. Google code : <http://googlecode.blogspot.com/2006/08/> announcing-tesseract- ocr.html (last accessed 8 January, 2012)
2. Smith, R. “An Overview of the Tesseract OCR” in proc. ICDAR 2007, Curitiba, Paraná, Brazil.
3. Open source Tesseract OCR Engine, available at <http://code.google.com/p/tesseract-ocr/>
4. Tesseract OCR training data downloads. Available at: <http://code.google.com/p/tesseract-ocr/downloads/list>
5. Tesseract-OCR An OCR Engine that was developed at HP Labs between 1985 and 1995... and now at Google.

Available at: <http://code.google.com/p/tesseract-ocr/wiki/TrainingTesseract3>

1. A fork of Tesseract Tools for Android. Available at: https://github.com/rmtheis/tess-two
2. ARCore, available at: <https://developers.google.com/ar>
3. ARCore — Google’s AR developer platform — provides simple yet powerful tool at: <https://arvr.google.com/arcore/>
4. OpenCV, available at: <https://opencv.org/>
5. OpenCV OCR and text recognition with Tesseract by Adrian Rosebrock at: [https://www.pyimagesearch.com/2018/09/17/opencv-ocr-and-text-recognition- with-tesseract/](https://www.pyimagesearch.com/2018/09/17/opencv-ocr-and-text-recognition-with-tesseract/)

[7] Deep Learning based Text Recognition (OCR) using Tesseract and OpenCV at [https://www.learnopencv.com/deep-learning-based-text-recognition-ocr-using- tesseract-and-opencv/](https://www.learnopencv.com/deep-learning-based-text-recognition-ocr-using-tesseract-and-opencv/)