

**ĐẠI HỌC QUỐC GIA TP. HỒ CHÍ MINH**

**TRƯỜNG ĐẠI HỌC CÔNG NGHỆ THÔNG TIN**

**KHOA CÔNG NGHỆ PHẦN MỀM**



# Báo cáo đồ án cuối kỳ

# [Scan, Edit and Save]

**Giảng viên hướng dẫn: Huỳnh Tuấn Anh**

**Sinh viên thực hiện:**

**Phan Anh Khoa - 17520645**

**Lớp: SE121.K21.PMCL**

**Môn học: Đồ Án 1**

**TP HCM, Ngày 6 tháng 7 năm 2020**

# 

**Mục lục**

**Chương 1**: Giới thiệu đề tài 4

**Chương 2**: Các công nghệ nền tảng

**2.1**: Android studio 7

**2.2**: OCR 7

**2.3**: Google APIs 8

**2.4**: Google Vision 8

**2.5**: Image Cropper 8

**2.6**: Text to Speech 9

**Chương 3**: Xây dựng ứng dụng

**3.1**: Tổng quan – Khảo sát các ứng dụng liên quan 10

**3.2**: Phân tích thiết kế phần mềm 11

**Chương 4**: Thử nghiệm – Đánh giá 20

**Chương 5**: Kết luận và hướng phát triển 21

**Phụ lục**: Tài liệu tham khảo 22

# CHƯƠNG 1: GIỚI THIỆU ĐỀ TÀI

Điện thoại di động nói chung, hay smartphone nói riêng là một trong những phát minh vĩ đại nhất của con người. Người người, nhà nhà, ai ai cũng sở hữu một chiếc smartphone cho riêng mình. Nhưng, smartphone ra đời để làm gì? Một và chỉ một mục đích: phục vụ con người, giúp cuộc sống con người thêm tiện nghi. Khả năng mang đi mọi lúc mọi nơi mà smartphone mang đến cho ta đã chứng minh ít nhiều điều đó. Trải qua bao nhiêu đời nâng cấp, cải tiến, mục đích chính của smartphone vẫn là làm sao để người sử dụng smartphone được tiện nghi, sử dụng hiệu quả và tối ưu chiếc smartphone của mình. Điều này không những phụ thuộc vào chiếc smartphone, mà còn là những phần mềm mà chiếc smartphone mang lại cho chúng ta – những gì mà chúng ta gọi là ứng dụng, là app.

Không những smartphone, một số công nghệ mới ra đời cũng để phục vụ cho các nhu cầu của con người, khi cuộc sống diễn ra ngày càng nhanh, tấp nập và bận bịu hơn. Optical Character Recognition (OCR) là một trong những ví dụ điển hình của công nghệ mới.

Optical Character Recognition (OCR) hay còn được gọi nhận diện kí tự quang học đã trở nên quen thuộc trong đời sống của chúng ta. Những phần mềm dung để scan text, biển số xe, hình ảnh, vv… là những ứng dụng tuyệt vời với tính tiện ích cao của công nghệ trên. Không chỉ trên những phần cứng chuyên dụng, OCR có thể được tích hợp vào những thiết bị di động thông minh, nâng cao tính cơ động và những lợi ích của công nghệ này trong những hoạt động thường ngày.

Vậy, chúng ta có một thiết bị bao gồm cả camera và bộ vi xử lí có thể mang đi mọi lúc mọi nơi. Và chúng ta có một công nghệ sử dụng hình ảnh, máy ảnh để scan chữ, văn bản. Nếu kết hợp chúng lại với nhau, ta sẽ có được gì nào?

Giả sử bạn đi gặp khách hàng. Khách hàng cho bạn xem tấm danh thiếp của họ, nhưng than ôi bạn lại không có mang bút và giấy để ghi lại vì bạn đã để quên chúng ở nhà. Khách hàng phải rời đi trong ít phút nữa, nhưng bạn lại phải làm phiền họ thêm vài phút để có thể viết thông tin cần thiết trên danh thiếp của khách hàng lên ứng dụng như Word hay Memo của điện thoại. Điều này khiến bạn mất điểm trong mắt khách hàng, và bạn ước có cách nào ghi lại thông tin danh thiếp của khách hàng chỉ qua một tấm ảnh hay một lần scan. Trường hợp tương tự khi bạn đi họp mà không muốn mang bút viết hay thậm chí không muốn dùng Word hay Memo trên điện thoại? Không cần phải lo, vì đã có Scan, Edit and Save (SES). Về cơ bản, SES là phần mềm sử dụng công nghệ OCR cho phép bạn trích xuất văn bản từ hình ảnh trong bộ nhớ hoặc trực tiếp hình ảnh từ camera, sau đó cho phép bạn chỉnh sửa, định dạng văn bản và cuối cùng là lưu văn bản nếu muốn. SES còn tích hợp thêm cả công nghệ Text To Speech cho phép mở rộng khả năng ứng dụng của app. Lợi ích của SES rất rộng. Bạn không mang bút viết mà chỉ mang chiếc smartphone của mình tới lớp? Không phải lo, SES sẽ giúp bạn lưu lại những gì giảng viên ghi lên bảng / slide khi bạn sử dụng khả năng OCR của nó. Bạn thấy một thông báo chỉ hiển thị trong vài phút, một câu viết nào đó hay hiển thị tức thời, hoặc đơn giản hơn là bạn đang vội mà không muốn bỏ thời gian ra ghi? Hãy sử dụng SES và để thời gian làm những việc quan trọng hơn. Và còn nữa, với khả năng OCR và lưu trữ kết quả đồng thời kết hợp Text to Speech, SES cũng là một công cụ tuyệt vời giúp bạn với việc học ngôn ngữ khi vừa lưu từ mới vừa cho phép bạn nghe cách phát âm của nó, hay chỉ đơn giản hơn, là khi bạn không thể phát âm chuẩn một từ nào đó trong quá trình học và quá lười để bật ứng dụng từ điển lên và tra từ đó. SES thích hợp để sử dụng cá nhân, bởi ứng dụng hướng đến sự tiện nghi của cá nhân khi giúp tiết kiệm thời gian cho việc đánh chữ và cho phép người dùng dành thời gian đó cho những công việc quan trọng hơn. Trong việc phát triển SES, nhóm phát triển hiện đang chú trọng đến công nghệ OCR, nghiên cứu về khả năng hậu xử lí đối với văn bản là kết quả của quá trình OCR đồng thời thêm những tính năng giúp mở rộng khả năng ứng dụng của SES đối với người sử dụng. Kết quả hướng tới của SES là giúp người sử dụng smartphone khai thác công nghệ OCR thông qua chính thiết bị của mình nhằm lấy và lưu lại văn bản mà không cần tốn thời gian đánh chữ trên thiết bị hoặc nhìn ảnh, từ đó giúp tiết kiệm thời gian và tăng độ chính xác, đặc biệt đối với văn bản số khi số chữ số quá nhiều và có khả năng gây rối mắt con người. Ngoài ra một số tính năng phụ như khả năng chỉnh sửa và định dạng lại văn bản cùng tính năng Text to Speech sẽ giúp người sử dụng tùy chỉnh văn bản và sử dụng theo mong muốn của họ.

CHƯƠNG 2: CÁC CÔNG NGHỆ NỀN TẢNG

**1. Android Studio:**

Android Studio là một môi trường phát triển tích hợp (IDE) chính thức cho hệ điều hành Android của Google, được xây dựng trên phần mềm IntelliJ IDEA của JetBrains và được thiết kế dành riêng cho phát triển Android. Nó có sẵn để tải xuống trên các hệ điều hành dựa trên Windows, macOS và Linux. Nó là sự thay thế cho Công cụ phát triển Android Eclipse (ADT) làm IDE chính cho phát triển ứng dụng Android gốc.

Đề án này sử dụng Android Studio và ngôn ngữ Java làm nền tảng để phát triển.

**2. OCR (Optical Character Recognition):**

Nhận dạng kí tự quang học là công nghệ cho phép người dùng chuyển đổi dữ liệu văn bản giấy thành dữ liệu kỹ thuật số. Được sử dụng rộng rãi như một hình thức nhập dữ liệu từ hồ sơ dữ liệu giấy in - cho dù tài liệu hộ chiếu, hóa đơn, báo cáo ngân hàng, biên lai máy tính, danh thiếp, thư, bản in dữ liệu tĩnh hoặc bất kỳ tài liệu phù hợp nào - đó là một phương pháp số hóa phổ biến được in các văn bản để chúng có thể được chỉnh sửa điện tử, tìm kiếm, lưu trữ gọn hơn, hiển thị trực tuyến và được sử dụng trong các quy trình của máy như điện toán nhận thức, dịch máy, (trích xuất) văn bản thành giọng nói, dữ liệu chính và khai thác văn bản.

**3. Google APIs:**

Là một bộ giao diện lập trình ứng dụng (API) do Google phát triển, cho phép giao tiếp với Google Services và sự tích hợp của chúng với các dịch vụ khác. Ví dụ về những điều này bao gồm Tìm kiếm, Gmail, Dịch hoặc Google Maps. Các ứng dụng của bên thứ ba có thể sử dụng các API này để tận dụng hoặc mở rộng chức năng của các dịch vụ hiện có.

Các API cung cấp chức năng như phân tích, máy học như một dịch vụ hoặc truy cập vào dữ liệu người dùng (khi được phép đọc dữ liệu được cung cấp). Một ví dụ quan trọng khác là bản đồ Google được nhúng trên một trang web, có thể đạt được bằng cách sử dụng API bản đồ tĩnh, API địa điểm hoặc API Google Earth.

**4. Google Vision:**

Google Vision API là một trong các APIs của google cung cấp chức năng phát hiện các đối tượng, khuôn mặt, văn bản in và viết tay từ hình ảnh bằng cách sử dụng các mô hình học máy được đào tạo trước. Bạn có thể tải từng hình ảnh lên công cụ và lấy nội dung của nó.



**5. Image Cropper:**

Là open-source library trên github. Được dùng để lựa chọn vị trí hình mà người dùng sẽ cắt để dễ dàng hơn trong việc chuyển đổi nội dung người dùng muốn sang dạng chữ.

**6. Text to Speech:**

Là một cộng nghệ hỗ trợ đọc to văn bản kỹ thuật số. Nó đôi khi được gọi là công nghệ đọc lớn tiếng. Text to Speech có thể lấy chữ trên máy tính hoặc thiết bị kỹ thuật số khác và chuyển đổi chúng thành âm thanh. TTS rất hữu ích cho những đứa trẻ phải vật lộn với việc đọc

CHƯƠNG 3: XÂY DỰNG ỨNG DỤNG

**1. Tổng quan – Khảo sát các ứng dụng liên quan:**

Hiện nay, công nghệ OCR trên mobile được áp dụng ngày càng rộng rãi. Nhìn chung, các ứng dụng được thiết kế để có thể sử dụng dữ liệu thu được sau quá trình scan để phục vụ một mục đích nhất định:

* Lưu trữ tài liệu ở nhũng định dạng text như .doc, .pdf,…vv.
* Lưu dữ liệu biển số xe phục vụ cho mục đích giữ xe hiệu quả hơn.
* Upload dữ liệu lên máy chủ xử lý và trả về kết quả (vd: Google Translate).

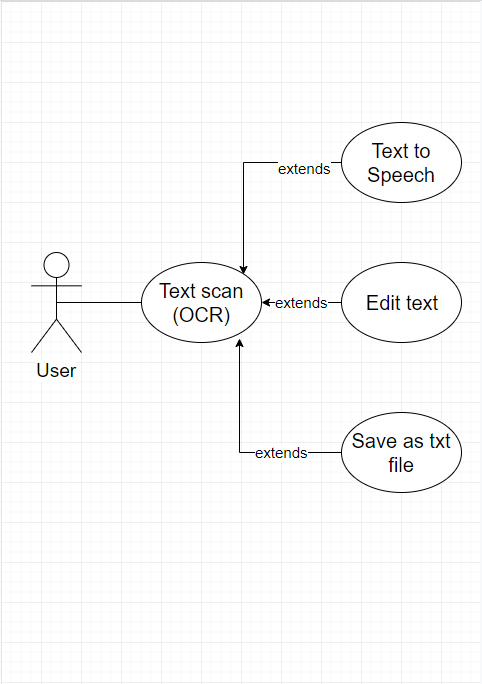
và còn một số ứng dụng phong phú khác.

Tuy nhiên, đa phần những phần mềm tốt đều tính phí cho toàn bộ phần mềm hoặc một số chức năng. Những phần mềm miễn phí thì phủ kín bởi các quảng cáo sau hoặc trước mỗi lần scan, gây phiền toái và bất tiện cho người sử dụng.

Dựa vào kết quả khảo sát trên, nhóm đặt ra mục tiêu phát triển một phần mềm text scan đơn giản, đáp ứng đủ các nhu cầu cơ bản của một phần mềm OCR (Scan – chỉnh sửa – lưu và xuất dữ liệu) dễ sử dụng và có khả năng phát triển cao hơn về sau.

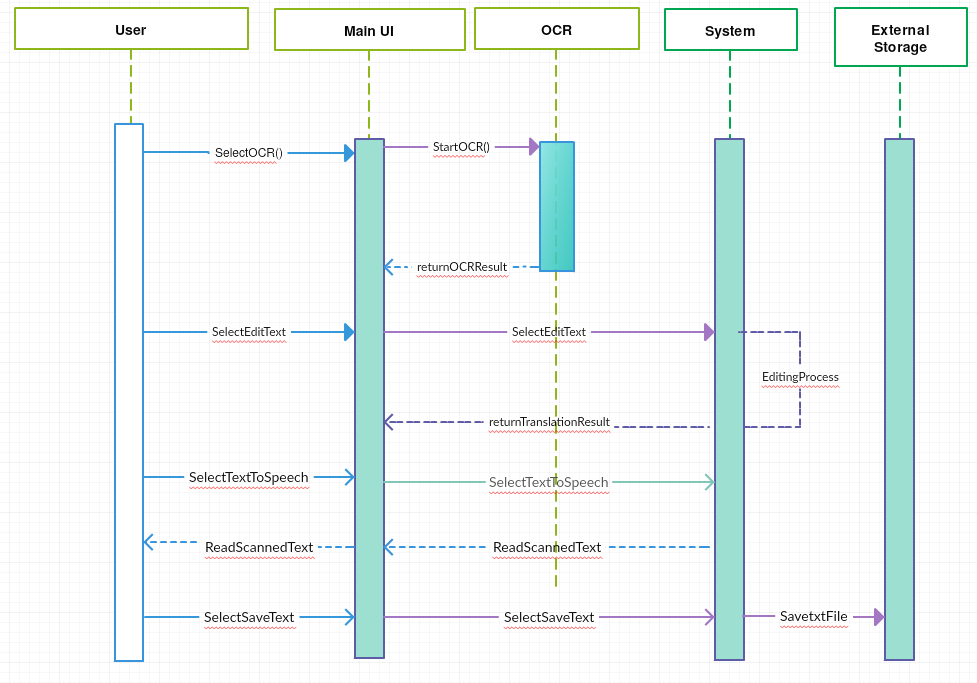
**2. Phân tích thiết kế phần mềm:**

**a/ Sơ đồ usecase:**



* Tác nhân: Chỉ có một tác nhân duy nhất là User (người sử dụng app)
* Các usecase:
  + **Text Scan**: Scan text từ hình ảnh chụp sẵn hoạc chụp ảnh mới.
  + **Text to Speech**: Đọc thành tiếng dữ liệu vừa được scan.
  + **Edit Text**: Chỉnh sửa nội dung của text được scan theo ý muốn của người dùng.
  + **Save As .txt File**: Lưu những thay đổi và xuất file (định dạng .txt).

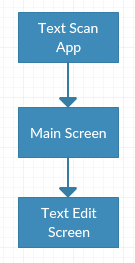
**b/ Biểu đồ tuần tự:**



Miêu tả trình tự hoạt động của ứng dụng: Người dùng sẽ chọn OCR thông qua hình ảnh đã chụp hoặc chụp ảnh mới, kết quả scan sẽ được trả về màn hình. Người dùng sau đó có thể yêu cầu ứng dụng đọc thành tiếng kết quả scan, chỉnh sửa tùy ý và sau đó lưu đoạn text vào một file exe trên bộ nhớ của điện thoại.

**c/ Thiết kế giao diện:**

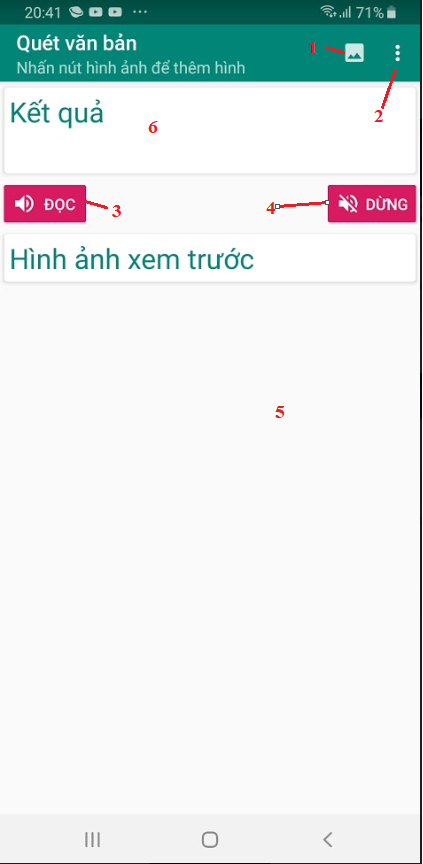
* **Sơ đồ liên kết màn hình:**



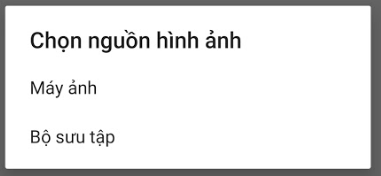
|  |  |  |
| --- | --- | --- |
| **STT** | **Màn hình** | **Chức năng** |
| 1 | Màn hình chính | Scan, text to speech, đổi ngôn ngữ, lưu và xuất txt. |
| 2 | Màn hình chỉnh sửa | Chỉnh sửa text đã được scan. |

* **Mô tả xử lý sự kiện từng màn hình:**

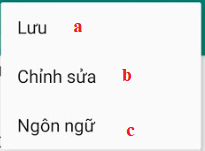
**Màn hình chính:**



1. Nút chọn ảnh: chọn ảnh để scan hoặc chụp ảnh mới

Lần đầu chạy chương trình sẽ có hộp thoại yêu cầu cấp phép truy cập camera và thư viện ảnh

1. Menu: bao gồm các chức năng Lưu, Chỉnh sửa và Đổi ngôn ngữ.



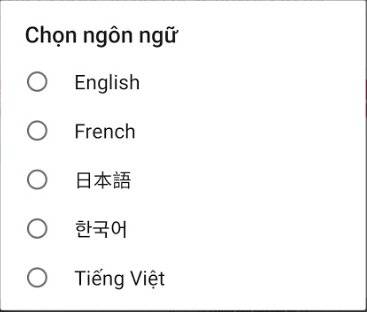
a/ Nút Lưu: Chọn chức năng này sẽ hiện lên thông báo



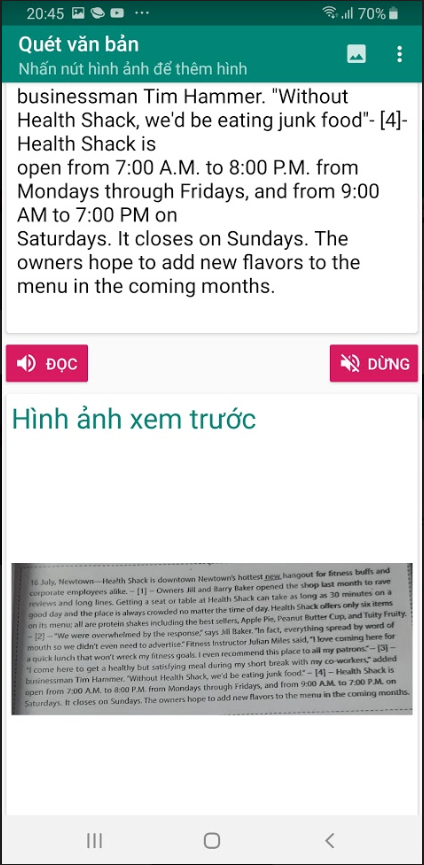
Đặt tên file và lưu

b/ Chỉnh sửa: chuyển sang màn hình chỉnh sửa.

c/ Ngôn ngữ: Cho phép đổi ngôn ngữ của ứng dụng sang các ngôn ngữ được hỗ trợ.

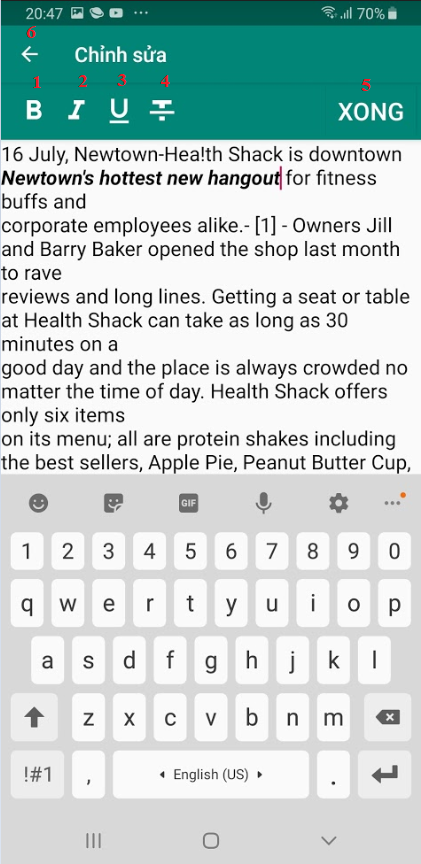


1. Nút Đọc: nhấn vào để đọc text đã được scan.
2. Nút Dừng: nhấn để dừng đọc đoạn text nếu đang đọc.
3. Image View: hiển thị hình ảnh được chọn hoặc vừa chụp.
4. Edit Text View: hiển thị kết quả scan được.



**Đây là màn hình kết quả sau một lần scan.**

**Màn hình chỉnh sửa:**



Bấm giữ và kéo lê con trỏ để chọn dòng chữ muốn chỉnh sửa

1. Nút Bold: In đậm chữ đã chọn.
2. Nút Italic: In nghiêng chữ đã chọn.
3. Nút Underline: Gạch dưới chữ đã chọn.
4. Nút Strike through: Gạch ngang chữ đã chọn.
5. Nút Hoàn tất: Lưu thay đổi và quay lại màn hình chính.
6. Nút Quay lại: Hủy thay đổi và quay lại màn hình chính.

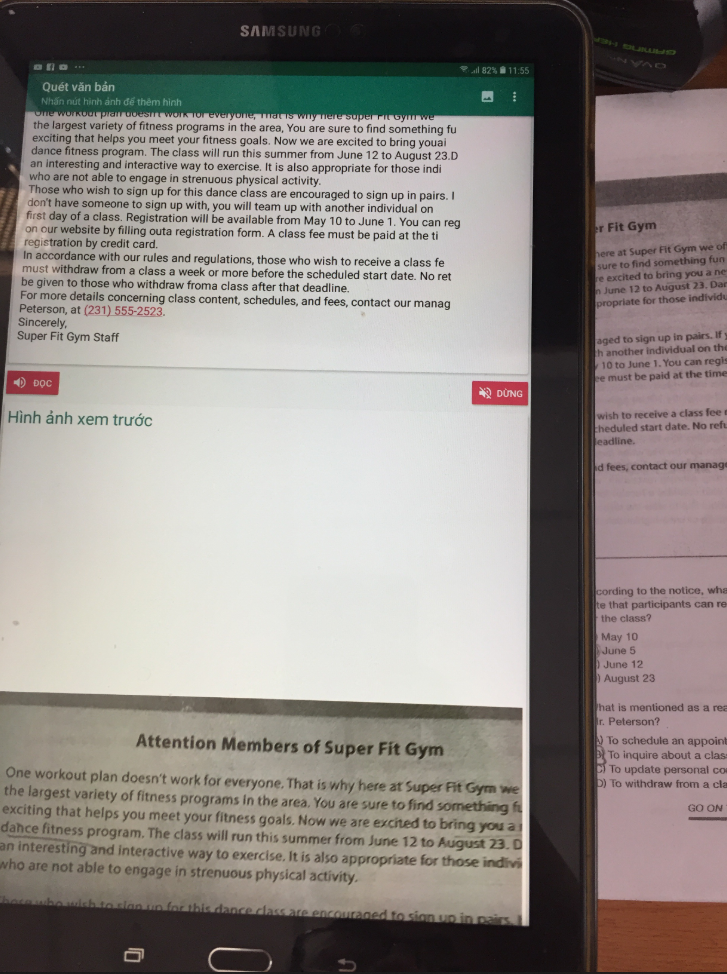
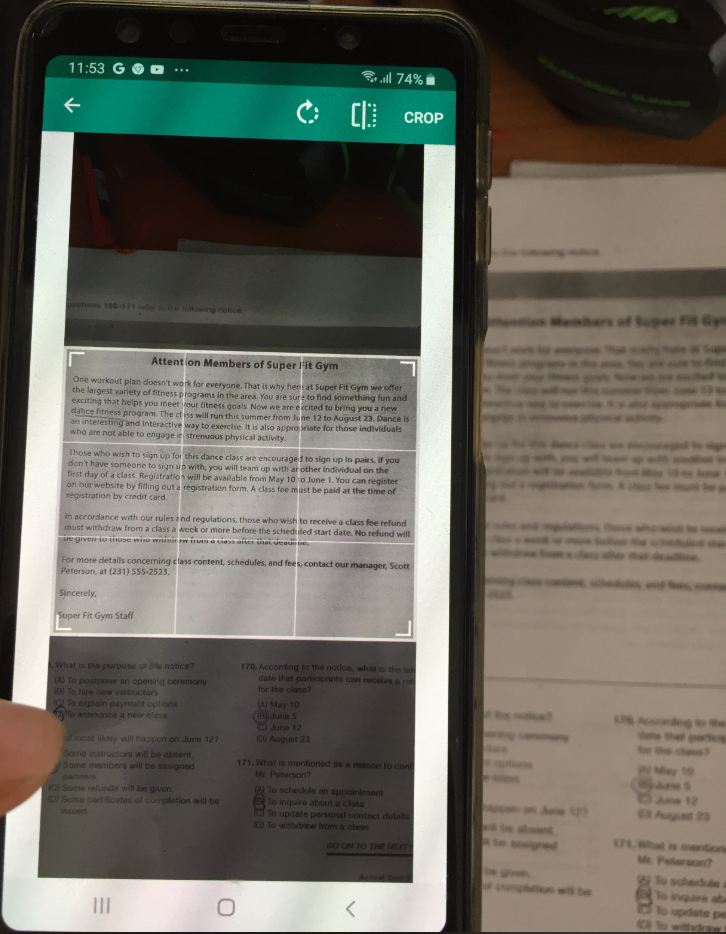
CHƯƠNG 4: THỬ NGHIỆM – ĐÁNH GIÁ

**1. Môi trường thử nghiệm:**

Ứng dụng được chạy thử trên các thiết bị hệ diều hành android với các phiên bản 8.0.0, phiên bản 8.1.0 và phiên bản 9.0.0

**2. Két quả thử nghiệm:**

Ứng dụng hoạt động bình thường trên các thiết bị nêu trên, chức năng hoạt động đầy đủ, không xuất hiện lỗi.

**3. Đánh giá kết quả sản phẩm:**

Sản phẩm có thể sử dụng để scan và lưu trữ được những thông tin, note, và những tài liệu khác trong hoạt dộng sống hằng ngày thông qua máy ảnh của smartphone.

Tuy nhiên, những tính năng hiện tại của sản phẩm chưa có tính cạnh tranh cao với những ứng dụng tương đồng khác trên thị trường, giao diện chưa thực sự cuốn hút.

CHƯƠNG 5: KẾT LUẬN VÀ HƯỚNG PHÁT TRIỂN

**1. Kết quả đạt được:**

Sau đề án này, tôi đã học thêm được rất nhiều kiến thức về lập trình ứng dụng di động nói chung và về công nghệ OCR nói riêng. Đây là những kiến thức cơ bản quan trọng và là nền tảng cho việc phát triển ứng dụng di động về sau.

**2. Hạn chế và hướng phát triển:**

Do công nghệ OCR là một công nghệ mới đối với tôi, nên sản phẩm đầu tiên còn nhiều mặt hạn chế:

* Chức năng của ứng dụng là những chức năng cơ bản của một ứng dụng OCR hoàn thiện, chưa thật sự nổi bật so với các sản phẩm trên thị trường.
* Giao diện sản phẩm còn đơn giản, thiếu sự cuốn hút.
* Khả năng định dạng văn bản không cao, thiếu một số chức năng như canh lề, chỉnh cỡ chữ, chỉnh font chữ, canh khoảng cách giữa các chữ, các dòng, etc.

Tuy nhiên sản phẩm này có tiềm năng và là nền tảng để phát triển những chức năng tiện lợi hơn và lôi cuốn hơn trong tương lai. Những hướng phát triển cho sản phẩm:

* Cải thiện giao diện sản phẩm.
* Tăng độ chính xác của OCR.
* Khả năng giữ nguyên layout format ban đầu của văn bản.
* Liên kết với kho lưu trữ đám mây để tiết kiệm dung lượng cho thiết bị di động.
* Khả năng chọn định dạng file output (hiện tại chỉ có thể xuất txt).

PHỤ LỤC: TÀI LIỆU THAM KHẢO

<https://github.com/ArthurHub/Android-Image-Cropper>

<https://developers.google.com/android/guides/releases>

<https://guides.codepath.com/android/Working-with-the-EditText>

<https://translate.google.com/?hl=vi>

<https://developer.android.com/reference/android/view/inputmethod/EditorInfo.html#IME_ACTION_SEARCH>

<https://stackoverflow.com/questions/46582552/how-do-i-search-for-a-word-in-a-textview>

<https://developer.android.com/reference/android/graphics/Typeface>

<https://developer.android.com/reference/android/text/SpannableString.html>

<https://developer.android.com/reference/android/widget/Button>

<https://developer.android.com/reference/android/widget/EditText>

<https://developer.android.com/reference/android/text/style/UnderlineSpan>

<https://developer.android.com/reference/android/graphics/Typeface.html#NORMAL>

<https://developer.android.com/reference/android/text/Selection.html#setSelection(android.text.Spannable,%20int,%20int)>

<https://developer.android.com/reference/java/lang/StringBuilder>

<https://github.com/lawloretienne/Trestle/issues/7>

<https://stackoverflow.com/questions/37824790/how-to-remove-underline-below-edittext-indicator/37825781>

<https://stackoverflow.com/questions/47797541/set-typeface-of-text-view-to-typeface-normal-doesnt-have-any-effect-why>

<https://developer.android.com/reference/android/text/style/StyleSpan.html#getStyle()>

<https://stackoverflow.com/questions/17921314/removing-style-from-selected-text-in-edittext?noredirect=1&lq=1>

<https://stackoverflow.com/questions/40963541/trying-to-bold-specific-text-in-android-studio-in-expandable-list>

<https://www.geeksforgeeks.org/android-how-to-send-data-from-one-activity-to-second-activity/>

<https://stackoverflow.com/questions/10979821/how-to-make-part-of-the-text-bold-in-android-at-runtime>

<https://www.youtube.com/watch?v=AD5qt7xoUU8>

<https://stackoverflow.com/questions/36307409/remove-style-on-spans>

<https://www.codota.com/code/java/methods/android.text.SpannableString/setSpan>

<https://stackoverflow.com/questions/10343838/get-value-of-span-text/38663735>

<https://stackoverflow.com/questions/17261207/remove-bold-from-textview-without-changing-other-attributes>

<https://www.youtube.com/results?search_query=android+studio+how+to+remove+bold>

<https://stackoverflow.com/questions/21569138/removing-italic-and-bold-html-tags-from-string/21569779>

<https://stackoverflow.com/questions/8471711/android-string-format-specify-bold>

<https://stackoverflow.com/questions/4792260/how-do-you-change-text-to-bold-in-android>

<https://stackoverflow.com/questions/4792260/how-do-you-change-text-to-bold-in-android>

<https://stackoverflow.com/questions/37696436/how-to-get-name-of-the-font-applied-on-the-textview>

<https://stackoverflow.com/questions/33535691/how-to-remove-span-style-from-spannablestring/33537780#33537780>

<https://stackoverflow.com/questions/29104552/in-android-studio-am-i-able-to-hide-and-show-a-label-textview-if-so-what-would>

<https://stackoverflow.com/questions/4297763/disabling-of-edittext-in-android>

<https://stackoverflow.com/questions/18012403/android-how-to-check-if-textview-is-bold-or-italic-programatically>

<https://stackoverflow.com/questions/29504749/how-to-check-that-the-text-is-bold-or-not?noredirect=1&lq=1>

<https://stackoverflow.com/questions/10714636/how-to-unbold-the-selected-text-in-edittext-android>

<https://stackoverflow.com/questions/3609174/android-insert-text-into-edittext-at-current-position>

<https://stackoverflow.com/questions/2163017/android-get-selection-of-text-from-edittext/28207750>

<https://www.guru99.com/string-length-method-java.html>

<https://examples.javacodegeeks.com/core-java/lang/string/java-string-length-example/>

<https://stackoverflow.com/questions/21668632/check-length-of-a-string-in-android>

<https://stackoverflow.com/questions/19924664/selecting-text-from-a-textview-or-edittext>

<https://stackoverflow.com/questions/1309629/how-to-change-colors-of-a-drawable-in-android>

<https://developer.android.com/reference/android/graphics/drawable/Drawable.html#setColorFilter%28int,%20android.graphics.PorterDuff.Mode%29>

<https://developer.android.com/reference/android/graphics/Paint.html#setColorFilter(android.graphics.ColorFilter)>

<https://developer.android.com/reference/android/graphics/ColorFilter.html>

<https://developer.android.com/reference/android/graphics/ColorFilter.html>

<https://developers.google.com/android/reference/com/google/android/gms/vision/text/TextBlock>

<https://stackoverflow.com/questions/21711468/center-a-drawable-in-a-button-with-background-and-no-text/21713921>

<https://stackoverflow.com/questions/10070762/how-to-get-selected-text-from-edittext-in-android>

<https://android--code.blogspot.com/2015/01/android-button-text-size-xml.html>

<https://developer.android.com/reference/android/widget/ImageButton>

<https://stackoverflow.com/questions/22966143/android-button-with-icon-and-text>

<https://stackoverflow.com/questions/37134750/add-items-to-android-toolbar>

<https://stackoverflow.com/questions/29025961/setsupportactionbar-toolbar-cannot-be-applied-to-android-widget-toolbar-error/39302444#39302444>

<https://medium.com/@suragch/adding-a-menu-to-the-toolbar-in-android-60d096f9fb89>

<https://stackoverflow.com/questions/35648913/how-to-set-menu-to-toolbar-in-android/35649219>

<https://www.youtube.com/watch?v=LxQDT5WdT3s>

<https://www.youtube.com/results?search_query=android+studio+how+to+add+item+to+toolbar>

<https://www.youtube.com/watch?v=DMkzIOLppf4>

<https://stackoverflow.com/questions/14597031/how-can-i-make-bold-and-underline-a-string-dynamically-in-android-code/14597164>

<https://stackoverflow.com/questions/4623508/how-to-set-the-font-style-to-bold-italic-and-underlined-in-an-android-textview>

<https://www.quora.com/How-can-I-make-any-selected-text-in-an-EditText-bold-or-italic-in-Android-Studio-and-store-it-into-an-SQLite-database>

<https://androidride.com/android-textview-bold/>

<https://www.tutorialspoint.com/how-to-make-a-specific-text-on-textview-bold-in-android>

<https://stackoverflow.com/questions/14371092/how-to-make-a-specific-text-on-textview-bold/14371107>

<https://www.codeproject.com/Questions/1113046/Bold-a-string-in-java-android-studio>

<https://stackoverflow.com/questions/4792260/how-do-you-change-text-to-bold-in-android>

<https://stackoverflow.com/questions/6148952/how-to-get-selected-text-and-value-android-listpreference>

<https://stackoverflow.com/questions/11078487/whats-toolscontext-in-android-layout-files>

<https://stackoverflow.com/questions/19452269/android-set-text-to-textview> ....