

TRƯỜNG ĐẠI HỌC CÔNG NGHỆ THÔNG TIN

KHOA CÔNG NGHỆ PHẦN MỀM



# BÁO CÁO ĐỒ ÁN MÔN HỌC

## Nhập môn ứng dụng di động

**Đề tài:** Ứng dụng từ điển

**GVLT:** Cô Phan Nguyệt Minh

**GVTH:** Thầy Huỳnh Tuấn Anh

**Lớp:** SE114.I13.PMCL

### DANH SÁCH NHÓM SINH VIÊN THỰC HIỆN

1. Hứa Vĩ Trung - 15520940
2. Nguyễn Đức Thông - 15520856
3. Nguyễn Hoàng Anh Tuấn - 15520971

## LỜI CẢM ƠN

Trên thực tế không có sự thành công nào mà không gắn liền với những sự hỗ trợ, giúp đỡ dù ít hay nhiều, dù trực tiếp hay gián tiếp của người khác. Trong suốt thời gian từ khi bắt đầu học tập ở giảng đường đại học đến nay, em đã nhận được rất nhiều sự quan tâm, giúp đỡ của quý thầy cô, gia đình và bạn bè. Với lòng biết ơn sâu sắc nhất, em xin gửi đến thầy cô Phan Nguyệt Minh, Hoàng Văn Hà, Huỳnh Tuấn Anh - Trường Đại Học Công Nghệ Thông Tin đã cùng với tri thức và tâm huyết của mình để truyền đạt vốn kiến thức quý báu cho chúng em trong suốt thời gian học tập tại trường. Và đặc biệt, trong học kỳ này, khoa đã tổ chức cho chúng em được tiếp cận với môn học mà theo em là rất hữu ích đối với sinh viên ngành khoa Công nghệ phần mềm cũng như tất cả các sinh viên thuộc các chuyên ngành khoa Công nghệ phần mềm khác. Đó là môn học "Nhập môn ứng dụng di động". Em xin chân thành cảm ơn thầy/cô đã tận tâm hướng dẫn chúng em qua từng buổi học trên lớp cũng như những buổi nói chuyện, thảo luận về lĩnh vực sáng tạo trong nghiên cứu khoa học. Nếu không có những lời hướng dẫn, dạy bảo của thầy thì em nghĩ sản phẩm này của em rất khó có thể hoàn thiện được. Một lần nữa, em xin chân thành cảm ơn thầy. Bài thu hoạch được thực hiện trong khoảng thời gian gần 3 tuần. Bước đầu đi vào thực tế, tìm hiểu về lĩnh vực sáng tạo trong nghiên cứu khoa học, kiến thức của em còn hạn chế và còn nhiều ngỡ ngàng. Do vậy, không tránh khỏi những thiếu sót là điều chắc chắn, em rất mong nhận được những ý kiến đóng góp quý báu của quý thầy cô và các bạn học cùng lớp để kiến thức của em trong lĩnh vực này được hoàn thiện hơn.

Lời cảm ơn chân thành nhất xin gửi đến quý thầy cô. Sau cùng, em xin kính chúc quý thầy cô trong trường thật dồi dào sức khỏe, niềm tin để tiếp tục thực hiện sứ mệnh cao đẹp của mình là truyền đạt kiến thức cho thế hệ mai sau

## LỜI MỞ ĐẦU

Ngày nay các loại sách báo, tư liệu cần được lưu trữ dưới dạng văn bản số rất phổ biến. Văn bản số có ưu điểm như cập nhật, sửa chữa, cũng như trao đổi nhanh chóng hơn so với văn bản in giấy truyền thống. Mặt khác, qua thời gian thì chất lượng văn bản in giấy sẽ kém đi nhưng văn bản số vẫn không bị hỏng. Từ đó, nảy sinh vấn đề làm cách nào để khôi phục lại những thông tin của sách báo dưới dạng văn bản số để có thể tái bản. Đây là một nhiệm vụ thực tế trong nhiều lĩnh vực, chẳng hạn như trong các thư viện và nhà xuất bản.

Có một số cách khác nhau để giải quyết bài toán chuyển đổi trên. Một biện pháp dễ thực hiện nhất là nhập lại nội dung của văn bản thông qua bàn phím. Mặc dù vậy, đây là một công việc thủ công trong thao tác chế bản nên nếu số lượng văn bản là quá lớn và mất nhiều thời gian sẽ dẫn tới nhiều sai sót. Giải pháp khác là tạo ra một chương trình nhận dạng văn bản tự động. Theo hướng này, sách báo được máy quét lưu trữ dưới dạng ảnh số, chương trình có chức năng nhận dạng ký tự và từ, từ đó chuyển đổi thành văn bản số.

Và đó là bài toán nhận dạng ký tự quang học (tiếng Anh: Optical Character Recognition, viết tắt là OCR) là quá trình rút trích ký tự trong ảnh cho ra dưới dạng văn bản để có thể lưu trữ, chỉnh sửa và tìm kiếm được. Lĩnh vực OCR là một nhánh trong ngành khoa học xử lý ảnh, tuy còn mới mẻ so với nhiều lĩnh vực khoa học khác nhưng nhanh chóng đã đạt được nhiều bước tiến quan trọng. Xuất phát từ nhu cầu thực tế là đưa các tài liệu lưu trữ trên giấy vào máy tính mà không phải đánh máy, nhiều công nghệ OCR ra đời (phát hành dưới dạng sản phẩm thương mại, miễn phí hoặc nguồn mở) đã và đang ứng dụng rộng rãi trong các lĩnh vực liên quan đến việc nhận dạng

Hiện nay công nghệ mạng internet phát triển mạnh mẽ, internet có mặt hầu hết khắp nơi trên thế giới. Công nghệ di động phát triển vượt bậc, việc sử dụng điện thoại bùng nổ, điện thoại di động trở thành mặt hàng bình dân mà ai cũng có thể sở hữu được. Nhiều công nghệ OCR tiên tiến được đưa ra dưới dạng mã nguồn mở và được các công ty, tổ chức lớn tài trợ phát triển. Với nền tảng công nghệ hiện có ở trên đã tạo ra nhiều giải pháp để xây dựng được nhiều hệ thống có khả năng đáp ứng cao với vấn đề đặt ra.

Nhận thấy tính cần thiết của công nghệ OCR, nhóm chúng em quyết định ứng dụng OCR trong phần mềm từ điển. Với những tính năng cơ bản cần có của một từ điển, Tri.t (tên phần mềm) của chúng em sử dụng công nghệ OCR để nhận dạng văn bản dưới hình ảnh được truyền vào camera. Nhằm phục vụ cho các mục đích dịch thuật hay cần phải tra nhanh một từ nào đó vô tình gặp. Với nền tảng android, công cụ hỗ trợ android studio và sự hỗ trợ của thầy cô, chúng em mong muốn tạo ra một phần mềm từ điển có ích và hiệu quả cho cộng đồng.

<b>LỜI CẢM ƠN</b> .....	2
<b>LỜI MỞ ĐẦU</b> .....	2
<b>CHƯƠNG I : GIỚI THIỆU ĐỀ TÀI</b> .....	4
I.1    GIỚI THIỆU .....	4
I.2    SỰ CẦN THIẾT CỦA ĐỀ TÀI.....	4
I.3    MỤC TIÊU ĐỀ TÀI .....	4
I.4    CÁC CHỨC NĂNG CHÍNH.....	5
<b>CHƯƠNG II: HƯỚNG DẪN ÁP DỤNG OCR TESSERACT VÀO ANDROID</b> .....	6
II.1    YÊU CẦU .....	6
II.2    TẠO DỰ ÁN ANDROID .....	6
II.3    KHAI BÁO DEPENDENCY .....	7
II.4    CHUẨN BỊ HÌNH ẢNH.....	8
II.5    THIẾT KẾ ACTIVITY .....	8
II.6    KHỞI TẠO CÁC PHƯƠNG THỨC CƠ BẢN .....	10
II.7    COPY TRAINING DATA VÀO THIẾT BỊ .....	11
II.8    XỬ LÝ HÌNH ẢNH .....	12
II.9    CHẠY ỨNG DỤNG.....	13
<b>CHƯƠNG III : PHÂN TÍCH YÊU CẦU</b> .....	14
III.1    DANH SÁCH CÁC YÊU CẦU .....	14
III.2    PHÂN TÍCH YÊU CẦU .....	15
<b>CHƯƠNG IV: THIẾT KẾ</b> .....	16
IV.1    CƠ SỞ DỮ LIỆU .....	16
IV.2    KIẾN TRÚC CHƯƠNG TRÌNH.....	19
IV.3    GIAO DIỆN .....	23
<b>CHƯƠNG V : CÀI ĐẶT VÀ THỬ NGHIỆM</b> .....	33
<b>CHƯƠNG VI : KẾT LUẬN</b> .....	34

# CHƯƠNG I : GIỚI THIỆU ĐỀ TÀI

## I.1 GIỚI THIỆU

Tên đề tài:

***"Xây dựng phần mềm ứng dụng từ điển tích hợp OCR Tesseract"***

Nhân cơ hội được học môn Nhập môn ứng dụng di động cùng với sự hướng dẫn tận tình của quý thầy cô Phan Nguyệt Minh, Hoàng Văn Hà, Huỳnh Tuấn Anh -Trường Đại Học Công Nghệ Thông Tin , chúng em quyết định làm phần mềm ứng dụng từ điển có tích hợp OCR Tesseract.

## I.2 SỰ CẦN THIẾT CỦA ĐỀ TÀI

1. Cho đến nay trên thế giới đã có rất nhiều phần mềm ứng dụng từ điển, nhưng đa số các ứng dụng đều giống nhau. Vậy với sự hỗ trợ của OCR, phần mềm của chúng em giúp mọi người dịch được văn bản trên ảnh.
2. Trong những năm gần đây cùng với sự bùng nổ của machine learning và computer vision, đây là thời cơ thích hợp nhất để ứng dụng công nghệ của machine learning vào các ứng dụng đã có mặt lâu đời. Sau khi tích hợp sẽ nâng cao giá trị của ứng dụng.
3. Nhận thấy sự khó khăn khi phải dịch những từ, văn bản trong sách vở , bìa , báo chí của mọi người.

Với những vấn đề được trình bày ở trên thì đây là thời điểm chín mùi để xây dựng ứng dụng từ điển sử dụng các công nghệ 4.0.

## I.3 MỤC TIÊU ĐỀ TÀI

Nhằm hỗ trợ mọi người có nhu cầu muốn dịch các từ câu trong sách vở, bìa báo một cách nhanh và tiện lợi nhất. Phá bỏ rào cản về mặt tâm lí bị ngắt mạch đọc khi phải dừng lại tra từ một từ không biết. Đây có lẽ là điểm nhấn lớn nhất khiến cho từ điển của chúng em khác với các phần mềm từ điển lâu đời khác.

Đối tượng nhắm tới :

- Sinh viên, học sinh những người hay tiếp cận với sách vở.
- Những người có thói quen đọc sách tiếng anh.
- Những người đi làm hay phải làm việc với tài liệu tiếng anh.

## I.4 CÁC CHỨC NĂNG CHÍNH

Chương trình gồm các chức năng chính của một từ điển thông thường cùng với tích hợp OCR nhận dạng kí tự quang học

Danh sách các chức năng :

1. Dịch Anh – Việt
2. Dịch Việt – Anh
3. Cửa sổ tra nhanh
4. Dịch văn bản
5. Sử dụng OCR
6. Từ yêu thích
7. Thiết lập

## CHƯƠNG II: HƯỚNG DẪN ÁP DỤNG OCR TESSERACT VÀO ANDROID

### II.1 YÊU CẦU

Để có thể sử dụng bài hướng dẫn này một cách dễ dàng, người đọc sẽ cần có một số phương tiện cơ bản sau:

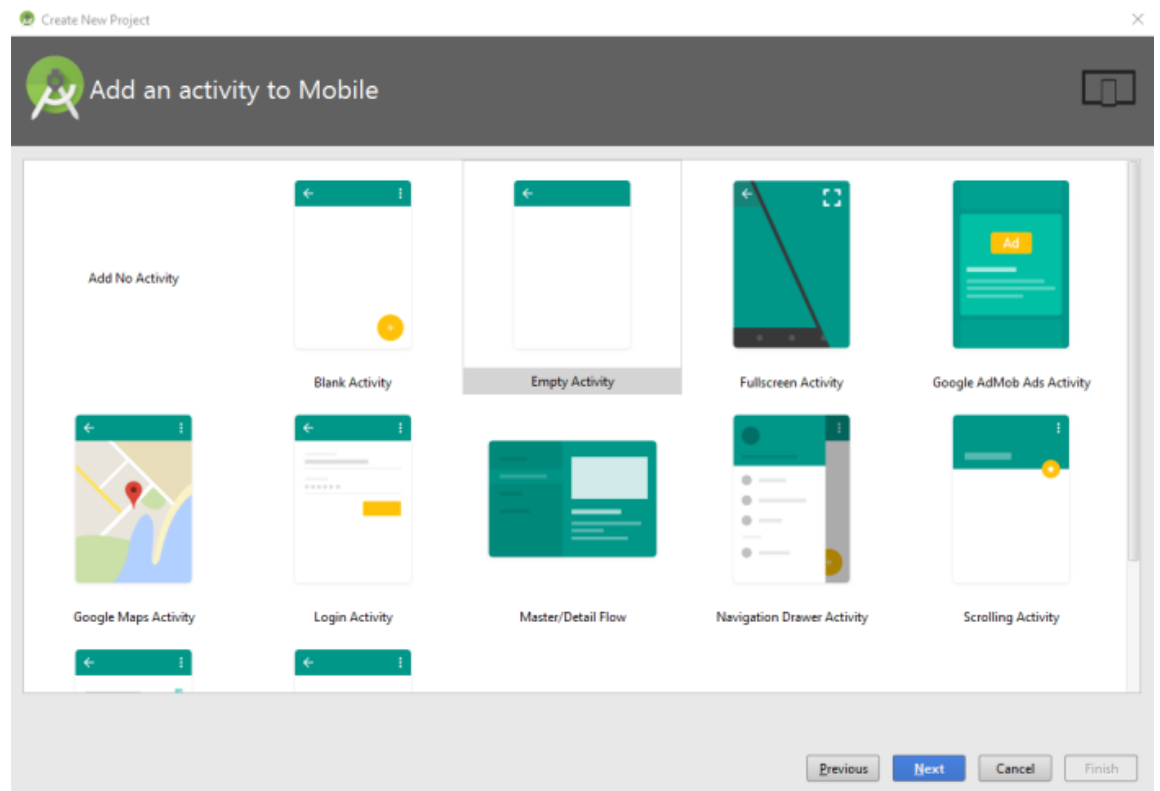
- Android Studio: IDE hỗ trợ lập trình Android
- Tập tin chứa dữ liệu ngôn ngữ
- Hình ảnh chữ

Bài hướng dẫn này sẽ gồm các bước chính như sau:

1. Tạo dự án Android
2. Khai báo dependency
3. Chuẩn bị hình ảnh
4. Thiết kế Activity
5. Khởi tạo các phương thức cơ bản
6. Copy training data vào thiết bị
7. Xử lý hình ảnh
8. Chạy ứng dụng

### II.2 TẠO DỰ ÁN ANDROID

Khởi động Android Studio, khởi tạo một dự án với khởi đầu là một



Empty Activity.

### II.3 KHAI BÁO DEPENDENCY

Chúng ta sẽ khai báo dependency cho ứng dụng trong tập tin **gradle.build** của thư mục **app** với nội dung như sau:

```
compile 'com.rmtheis:tess-two:5.4.1'
```

Như vậy, tập tin gradle.build của chúng ta sẽ có nội dung phần dependency như sau:

```
dependencies {  
  
    compile fileTree(dir: 'libs', include: ['*.jar'])  
  
    androidTestCompile('com.android.support.test.espresso:espresso-  
core:2.2.2', {  
        exclude group: 'com.android.support', module: 'support-annotations'  
    })  
}
```



```
})  
compile 'com.android.support:appcompat-v7:25.3.1'  
compile 'com.rmtheis:tess-two:5.4.1'  
testCompile 'junit:junit:4.12'}
```

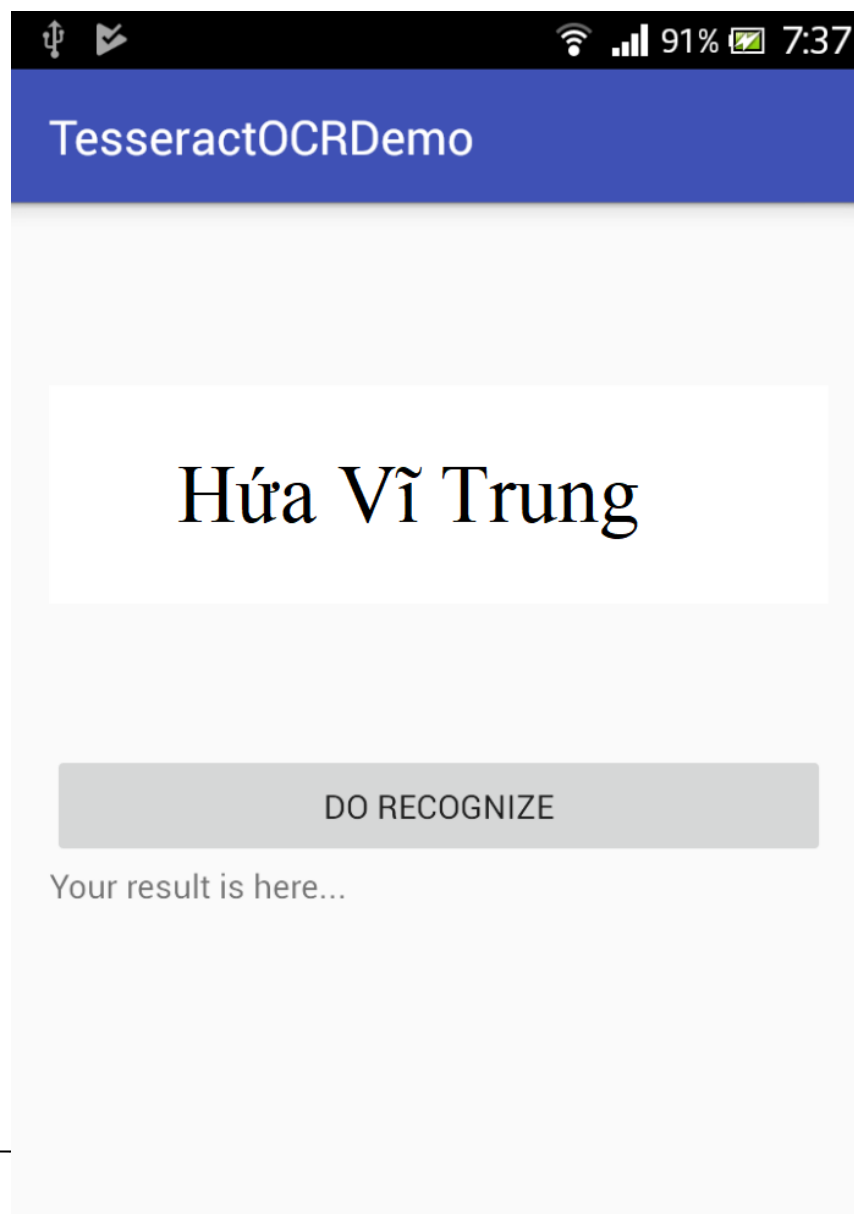
## II.4 CHUẨN BỊ HÌNH ẢNH

Với một ứng dụng OCR đơn giản, ta nên chuẩn bị một hình ảnh đơn giản với chữ đen, rõ ràng trên nền trắng

# Hứa Vĩ Trung

## II.5 THIẾT KẾ ACTIVITY

Trong bài hướng dẫn này, chúng ta sẽ xây dựng ứng dụng Android



nhận diện văn bản đơn giản gồm 01 ImageView, 01 button và 01 TextView

Như vậy, giao diện trên có thể sẽ được định nghĩa bởi file layout như sau

```
<ImageView

    android:id="@+id/img_input"

    android:layout_width="match_parent"

    android:layout_height="wrap_content" />


<Button

    android:id="@+id/btn_rec"

    android:layout_width="match_parent"

    android:layout_height="wrap_content"

    android:onClick="doRecognize"

    android:text="@string/btn_action_name"

    android:layout_below="@id/img_input"/>


<TextView

    android:id="@+id/txt_result"

    android:layout_width="wrap_content"

    android:layout_height="wrap_content"

    android:text="Your result is here..."

    android:layout_below="@id/btn_rec"/>
```

## II.6 KHỞI TẠO CÁC PHƯƠNG THỨC CƠ BẢN

Trước hết, bên cạnh các phương thức cơ bản để hiển thị một activity trong ứng dụng Android, bạn sẽ cần phải khởi tạo một biến thuộc class TessBaseAPI để sử dụng các chức năng liên quan đến Tesseract OCR. Ở đây, mình sẽ thực hiện như sau:

```
public class MainActivity extends AppCompatActivity {

    private TessBaseAPI m_tess;

    @Override

    protected void onCreate(Bundle savedInstanceState) {

        super.onCreate(savedInstanceState);

        setContentView(R.layout.activity_main);

        initImageView();

        try {

            prepareLanguageDir();

            m_tess = new TessBaseAPI();

            m_tess.init(getFilesDir(), "vie");
        } catch (Exception e) {
            // Logging here
        }
    }

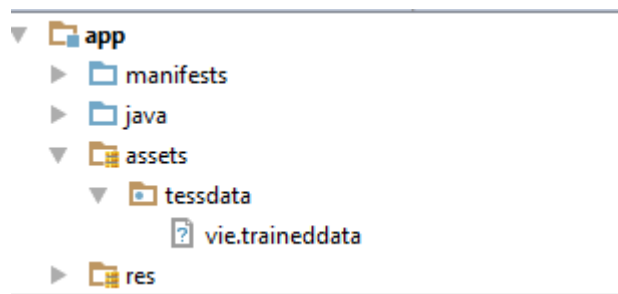
    private void initImageView() {
        ImageView imgView = (ImageView)
        findViewById(R.id.img_input);
        Bitmap input = BitmapFactory.decodeResource(getResources(),
        R.drawable.vie_tnq);
        imgView.setImageBitmap(input);
    }
}
```

Ta có thể thấy được các API của Tesseract sẽ được truy cập thông qua biến `m_tess` (class `TessBaseAPI`) và để khởi tạo biến này, ta sẽ cần chỉ định đường dẫn của thư mục chứa tập tin dữ liệu ngôn ngữ và loại ngôn ngữ được sử dụng. Ở đây, chúng ta sẽ dùng thư mục bên trong thư mục `data` của ứng dụng và ngôn ngữ là tiếng Việt.

Nên lưu ý rằng, thư mục được chỉ định phải chứa thư mục con có tên `tessdata`. Thư mục `tessdata` đó sẽ chứa các tập tin dữ liệu ngôn ngữ.

## II.7 COPY TRAINING DATA VÀO THIẾT BỊ

Trước tiên, chúng ta cần phải chuẩn bị tập tin dữ liệu ngôn ngữ. Bạn có thể download dữ liệu cần thiết được chia sẻ trên [Github tessdata](#). Ở đây, mình sẽ download tập tin **vie.traineddata** và lưu trong thư mục **assets/tessdata** của ứng dụng.



Như đã nói, chúng ta sẽ lưu tập tin trong thư mục **assets** nên ứng dụng sẽ không thể truy cập trên thiết bị. Do đó, chúng ta sẽ cần có một bước chuyển file vào thiết bị thông qua phương thức **prepareLanguageDir()** như sau:

```
public class MainActivity extends AppCompatActivity {

    // Other methods...

    // copy file from assets to another folder due to accessible

    private void copyFile() throws IOException {

        // work with assets folder

        AssetManager assMng = getAssets();

        InputStream is = assMng.open("tessdata/vie.traineddata");
        OutputStream os = new FileOutputStream(getFilesDir() +
            "/tessdata/vie.traineddata");
```

```

byte[] buffer = new byte[1024];
int read;
while ((read = is.read(buffer)) != -1) {
    os.write(buffer, 0, read);
}

is.close();
os.flush();
os.close();
}

private void prepareLanguageDir() throws IOException {
    File dir = new File(getFilesDir() + "/tessdata");
    if (!dir.exists()) {
        dir.mkdirs();
    }

    File trainedData = new File(getFilesDir() +
"/tessdata/vie.traineddata");
    if (!trainedData.exists()) {
        copyFile();
    }
}

```

## II.8 XỬ LÝ HÌNH ẢNH

Tới bước này, chúng ta đã hoàn tất những điều kiện cơ bản để có thể nhận diện ký tự trên một hình ảnh rõ ràng. Tiếp theo, chúng ta sẽ sử dụng API của Tesseract để nhận diện các ký tự. Với bài hướng dẫn này, việc nhận diện sẽ được khởi động bởi sự kiện nhấn button “Do Recognize” như sau:

```

public class MainActivity extends AppCompatActivity {

    // Other methods ...

    public void doRecognize(View view) {

        if (m_tess == null) {

            return;

        }

        try {

```

```
m_tess.setImage(BitmapFactory.decodeResource(getResources(),
R.drawable.vie_tnq));

String result = m_tess.getUTF8Text();

TextView resultView = (TextView)
findViewById(R.id.txt_result);

resultView.setText(result);

} catch (Exception e) {

    // Do what you like here...

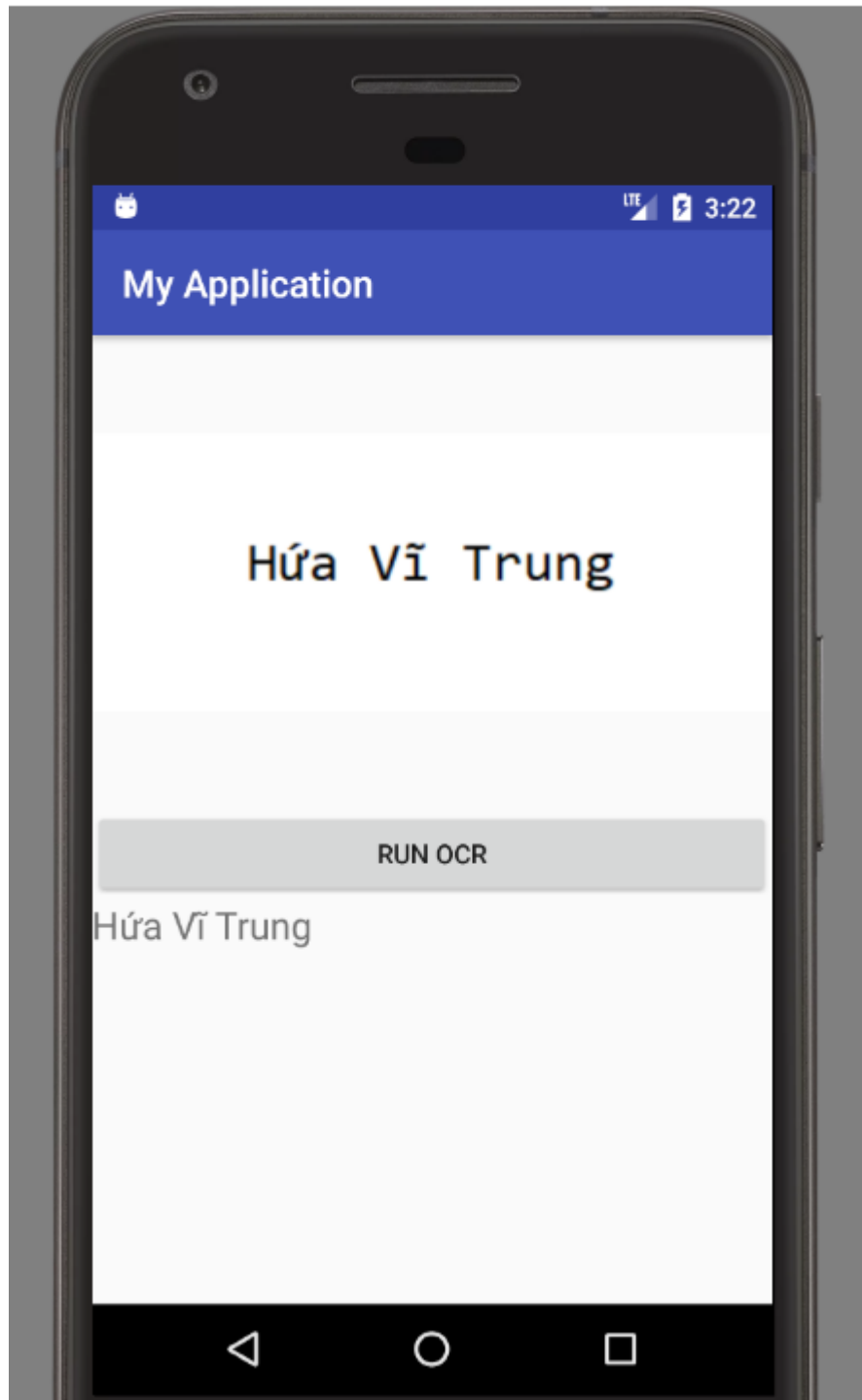
}

}

}
```

## II.9 CHẠY ỨNG DỤNG

Và bây giờ, bạn chạy thử ứng dụng và xem xem có đạt được kết quả mong muốn



## CHƯƠNG III : PHÂN TÍCH YÊU CẦU

### III.1 DANH SÁCH CÁC YÊU CẦU

STT	Tên yêu cầu	Ghi chú
1	Tra từ điển Anh Việt	
2	Tra từ điển Việt Anh	
3	Dịch văn bản	

4	Cửa sổ tra nhanh	
5	Ứng dụng OCR	
6	Lập mục từ yêu thích	
7	Thay đổi thiết lập	

### III.2 PHÂN TÍCH YÊU CẦU

Bảng trách nhiệm yêu cầu:

STT	Yêu cầu	Người dùng	Phần Mềm	Ghi chú
1	Tra từ điển Anh Việt	Cung cấp từ tiếng anh cần dịch	Kiểm tra từ nhập vào với cơ sở dữ liệu, nếu có thì trả về kết quả	Cho phép xem, sửa, xóa
2	Tra từ điển Việt Anh	Cung cấp từ tiếng việt cần dịch	Kiểm tra từ nhập vào với cơ sở dữ liệu, nếu có thì trả về kết quả	Cho phép xem, sửa, xóa
3	Dịch văn bản	Cung cấp văn bản tiếng anh cần dịch	Lấy văn bản nhập vào dịch trên google translate và trả về kết quả cho người dùng	Cho phép xem, sửa, xóa
4	Cửa sổ tra nhanh	Cung cấp một cách thứ	Thu nhỏ cửa sổ tiện lợi cho việc tra nhanh	Cho phép xem, kích hoạt
5	Ứng dụng OCR	Chụp văn bản hay từ cần dịch	Nhận diện hình ảnh chuyển sang văn bản và tiến hành dịch	Cho phép xem, kích hoạt và chụp hình
6	Lập mục từ yêu thích	Lưu từ yêu thích, từ đặc biệt, từ muốc được lưu ý	Lưu danh sách từ yêu thích của người dùng và tạo noti để thông báo trên điện thoại	Cho phép xem, thêm, sửa , xóa thông tin
7	Thay đổi thiết lập	Thay đổi ngôn ngữ, chế độ thêm và xem thông tin phần mềm	Thay đổi theo sự lựa chọn của người dùng	Cho phép xem, sửa.



## CHƯƠNG IV: THIẾT KẾ

### IV.1 CƠ SỞ DỮ LIỆU

#### 1. Hệ quản trị cơ sở dữ liệu sqlite

Nói một cách đơn giản SQLite là phần mềm quản lý cơ sở dữ liệu (DBMS) tương tự như Mysql, PostgreSQL... Đặc điểm của SQLite là gọn, nhẹ, đơn giản. Chương trình gồm 1 file duy nhất vốn vẹn chưa đến 400kB, không cần cài đặt, không cần cấu hình hay khởi động mà có thể sử dụng ngay. Dữ liệu Database cũng được lưu ở một file duy nhất. Không có khái niệm user, password hay quyền hạn trong SQLite Database.

SQLite không thích hợp với những hệ thống lớn nhưng ở quy mô vừa tầm thì SQLite phát huy uy lực và không hề yếu kém về mặt chức năng hay tốc độ. Với các đặc điểm trên SQLite được sử dụng nhiều trong việc phát triển, thử nghiệm ... và là sự lựa chọn phù hợp cho những người bắt đầu học Database.

SQLite Engine không là một Standalone Process giống như các cơ sở dữ liệu khác, bạn có thể liên kết nó một cách tĩnh hoặc một cách động tùy theo yêu cầu với ứng dụng của bạn. SQLite truy cập các file lưu giữ của nó một cách trực tiếp.

```
@00-database-info
- This is the English-Vietnamese dictionary database of the Free Vietnamese Dictionary Project. It contains more than 109.000 e
- This database was compiled by Ho Ngoc Duc and other members of the Free Vietnamese Dictionary Project (http://www.informatik.
- About 5.000 technical terms were imported from VACETS Dictionary Project (http://www.saigon.com/~diction/)
- Copyright (C) 1997-2003 The Free Vietnamese Dictionary Project
- Last updated 2004-07-01
- This program is free software; you can redistribute it and/or modify it under the terms of the GNU General Public License as ;
- This program is distributed in the hope that it will be useful, but WITHOUT ANY WARRANTY
- See the GNU General Public License for more details.

@00-database-short
- FVDP English-Vietnamese dictionary

@00-database-url
- http://www.informatik.uni-leipzig.de/~duc/Dict/
@a /ei, e/
* danh từ, số nhiều as, a's
- (thông tục) loại a, hạng nhất, hạng tốt nhất hạng rất tốt
=his health is a+ sức khoẻ anh ta vào loại a
- (âm nhạc) la
=a sharp+ la thăng
=a flat+ la giáng
- người giả định thứ nhất; trường hợp giả định thứ nhất
=from a to z+ từ đầu đến cuối, tường tận
=not to know a from b+ không biết tí gì cả; một chữ bẻ đôi cũng không biết
* mạo từ
- một; một (nhu kiểu); một (nào đó)
=a very cold day+ một ngày rất lạnh
=a dozen+ một tá
=a few+ một ít
=all of a size+ tất cả cùng một cỡ
=a Shakespeare+ một (văn hào nhu kiểu) Sêch-xpia
=a Mr Nam+ một ông Nam (nào đó)
- cái, con, chiếc, cuốn, người, đứa...;
=a cup+ cái chén
=a knife+ con dao
```

## 2. Xây dựng cơ sở dữ liệu

Bước 1 : Kiểm định dữ liệu từ file text – file dữ liệu lấy từ nguồn miễn phí

 <b>English</b>	2009-08-05	37	
 <b>Italian</b>	2008-01-22	0	
 <b>Spanish</b>	2007-07-06	0	
 <b>French</b>	2007-07-06	0	
 <b>Japanese</b>	2007-07-06	3	
 <b>Russian</b>	2007-07-06	2	
 <b>Korean</b>	2007-07-06	0	
 <b>German</b>	2007-07-06	0	
 <b>Chinese</b>	2007-07-06	1	
 <b>Czech</b>	2007-07-06	0	
 <b>Norwegian</b>	2007-07-06	0	
 <b>Portuguese</b>	2007-07-06	0	

Bước 2 : Đọc dữ liệu từ file text để tạo database. Sử dụng sqlite để quản lý các thao tác nhập, xóa, sửa, tìm kiếm.

Dùng hàm loadWord() trong file DictionaryDatabase.java để đọc dữ liệu từ file anhviet109k.txt, trong đó dùng hàm readLine() để đọc từng dòng

WORD	DEFINITION
a	ei, ə\$* danh từ, số nhiều as, a's\$- (thông tục) loại a, hạng nhất, hạng tốt nhất hạng rấ...
a b c	'eibi:'si:\$* danh từ\$- bảng chữ cái\$- khái niệm cơ sở, cơ sở\$a_b_c of chemistry+ kh...
a b c - book	'eibi:'si:buk\$* danh từ\$- sách vở lòng, sách học văn\$
a font	ɑ:'fɔ:n\$* phó từ\$- thấu triết, cận kề, rõ ngọn ngành\$
a fortiori	'ei,fo:'ti:'ɔ:rai\$* phó từ\$- huống hồ, huống là\$
a la carte	'ɑ:la:'ka:t\$* phó từ\$- theo món, gọi theo món, đặt theo món\$=to dine a_la_carte+ ăn ...
a la mode	,ɑ:lə'moud\$* phó từ\$- hợp thời trang, đúng mốt\$
a posteriori	'eipɒs,teri:'ɔ:rai\$* phó từ & tính từ\$- theo phép quy nạp\$=method a_posteriori+ phươn...
a priori	'eiprai:'ɔ:rai\$* danh từ & phó từ\$- theo cách suy diễn, theo cách diễn dịch\$- tiên nghiệ...
a-bomb	'ei'bɒm\$* danh từ\$- bom nguyên tử\$
a-going	ə'gəʊɪŋ\$* tính từ & phó từ\$- đang chạy, đang chuyển động; đang hoạt động, đang tiế...
a-plenty	ə'plenti\$* phó từ\$- (từ Mỹ,nghĩa Mỹ) nhiều, dồi dào, phong phú\$
a-power	'ei'paʊə\$* danh từ\$- năng lượng nguyên tử\$- cường quốc nguyên tử\$
a.d.	'ei'di:\$* (viết tắt) của Anno Domin\$- sau công nguyên\$=1540 a.d.+ năm 1540 sau côn...
a.m.	'ei'em\$* phó từ\$- (xem) ante_meridiem\$
aard-wolf	'ɑ:d,wʊlf\$* danh từ\$- (động vật học) chó sói đất (Nam Phi)\$
aasvogel	'ɑ:sfəʊgəl\$* danh từ\$- (động vật học) con kền kền (Nam Phi)\$
aba	'ɑ:bə\$* danh từ\$- áo aba (áo ngoài giống hình cái túi người A-Rập)\$
abaci	'æbəkes\$* danh từ, số nhiều abaci, abacuses\$- bàn tính\$=to move counters of an ab...
aback	ə'baek\$* phó từ\$- lùi lại, trở lại phía sau\$=to stand aback from+ đứng lùi lại để tránh\$...

trong file text và gộp lại thành một dòng definition trong database.

Bước 3 : Đọc dữ liệu từ database và đưa vào sử dụng trong chương trình bằng hàm query() đã được hỗ trợ bởi sqlite. Nhờ đó chúng ta có thể truy xuất vào

```

@Nullable
private Cursor query(String selection, String[] selectionArgs, String[] columns, String query) {
    SQLiteQueryBuilder builder = new SQLiteQueryBuilder();
    String table = String.valueOf(query.charAt(0));
    builder.setTables(table);

    Cursor cursor = builder.query(mDatabaseOpenHelper.getReadableDatabase(),
        columns, selection, selectionArgs, null, null, null);

    if (cursor == null) {
        return null;
    } else if (!cursor.moveToFirst()) {
        cursor.close();
        return null;
    }
    return cursor;
}

```

database một cách dễ dàng nhất có thể.

### 3. Một vài thao tác trên cơ sở dữ liệu

```
query(SQLiteDatabase db, String[] projectionIn, String selection, String[] selectionArgs, String groupBy, String having, String sortOrder)
```

đây là câu lệnh chính của chương trình dung để thực thi sqlite câu lệnh này sẽ trả về cursor. Biến cursor này sẽ nối tất cả các biến lại để tạo thành một câu query để truy vấn xuống database để lấy dữ liệu

## IV.2 KIẾN TRÚC CHƯƠNG TRÌNH

### IV.3 **public class** camera2 **extends** AppCompatActivity

Tên biến	String <b>TAG</b>
	Button <b>takePictureButton</b>
	TextureView <b>textureView</b>
	TextView <b>textViewW</b>
	TextView <b>textViewH</b>
	RectangleView <b>rectangleView</b>
	<b>final</b> SparseIntArray <b>ORIENTATIONS</b>
	String <b>cameraId</b>
	CameraDevice <b>cameraDevice</b>
	CameraCaptureSession <b>cameraCaptureSessions</b>
	CaptureRequest <b>captureRequest</b>
	CaptureRequest.Builder <b>captureRequestBuilder</b>
	Size <b>imageDimension</b>
	File <b>file</b>
	<b>final int</b> <b>REQUEST_CAMERA_PERMISSION</b>
	<b>boolean</b> <b>mFlashSupported</b>
	Handler <b>mBackgroundHandler</b>
	HandlerThread <b>mBackgroundThread</b>
Tên hàm	<b>public void</b> onWindowFocusChanged( <b>boolean</b> hasFocus)
	<b>protected void</b> startBackgroundThread()
	<b>protected void</b> stopBackgroundThread()
	<b>protected void</b> takePicture()
	<b>protected void</b> createCameraPreview()
	<b>private void</b> openCamera()
	<b>protected void</b> updatePreview()
	<b>private void</b> closeCamera()
	<b>public void</b> onRequestPermissionsResult( <b>int</b> requestCode, <b>@NonNull</b> String[] permissions, <b>@NonNull</b> <b>int</b> [] grantResults)
	<b>protected void</b> onResume()
	<b>protected void</b> onPause()
	<b>protected void</b> startBackgroundThread()

### IV.4 **public class** FloatingViewService **extends** Service

Tên biến	<b>private</b> WindowManager <b>mWindowManager</b>
	<b>private</b> View <b>mFloatingView</b>
	SearchView <b>searchViewInChatHead</b>
Tên hàm	<b>public</b> FloatingViewService()
	<b>public</b> IBinder onBind(Intent intent)
	<b>public void</b> onCreate()
	<b>private boolean</b> isViewCollapsed()

#### IV.5 **public class** LanguagesSetting **extends** AppCompatActivity

Tên biến	Button <b>btnSettingEnglish</b>
	Button <b>btnSettingGerman</b>
	Button <b>btnSettingVietnamese</b>
Tên hàm	<b>protected void</b> onCreate(Bundle savedInstanceState)

#### IV.6 **public class** ListViewAdapter **extends** BaseAdapter

Tên biến	Context <b>mContext</b>
	LayoutInflater <b>inflater</b>
	<b>private</b> List<String> <b>listAllWord</b>
	<b>private</b> ArrayList<String> <b>arraylist</b>
Tên hàm	<b>public</b> ListViewAdapter(Context context, List<String> wordList)
	<b>public class</b> ViewHolder
	<b>public int</b> getCount()
	<b>public</b> String getItem( <b>int</b> position)
	<b>public long</b> getItemId( <b>int</b> position)
	<b>public</b> View getView( <b>final int</b> position, View view, ViewGroup parent)
	<b>public void</b> filter(String charText)

#### IV.7 **public class** FavoriteWordActivity **extends** AppCompatActivity

Tên biến	LinearLayout <b>linearLayoutFavorieWord</b>
	DictionaryDatabase <b>dictionaryDatabase</b>
	<b>int</b> <b>NotificationID</b>
Tên hàm	<b>protected void</b> onCreate(@Nullable Bundle savedInstanceState)
	<b>public void</b> sendNotification(View view, String word)
	<b>void</b> startTranslate(String itemValue)

#### IV.8 **public class** MainActivity **extends** LocalizationActivity

Tên biến	SearchView SearchWord
	ImageView invSetting
	ListView lsvWord
	ListView lsvHistory
	ListViewAdapter adapter

	ListViewAdapter[] adapterlist
	ListViewHistoryAdapter adapterHistoryList
	ArrayList<String> listHistoryWord
	public ArrayList<String>[]
	String definition
	final int a
	Button btnEngVieDict
	Button btnFloatingWidget
	Button btnFavoriteWord
	Button btnSetting
	Button btnTranslateText
Tên hàm	<b>public void</b> onResume()
	<b>private void</b> initializeView()
	<b>void</b> startTranslate(String itemValue)
	<b>public boolean</b> onQueryTextSubmit(String query)
	<b>public boolean</b> onQueryTextChange(String newText)
	<b>public void</b> onActivityResult( <b>int</b> requestCode, <b>int</b> resultCode, Intent data)
	<b>public void</b> onBackPressed()

#### IV.9 **public class** OCR **extends** AppCompatActivity

Tên biến	Button btnRunOCR
	final int OCR_RESULT
	String mCurrentPhotoPath
	private TessBaseAPI mTess
	String datapath
	ImageView imvPicture
Tên hàm	<b>private void</b> copyFiles(String lang)
	<b>private void</b> checkFile(File dir, String lang)
	<b>protected void</b> onActivityResult( <b>int</b> requestCode, <b>int</b> resultCode, Intent data)

#### IV.10 **public class** RectangleView **extends** View

Tên biến	private Paint mPaint
	public Point point1
	public Point point2
	public Point point3
	public Point point4
Tên hàm	<b>public</b> RectangleView(Context context)
	<b>public</b> RectangleView(Context context, AttributeSet attrs)
	<b>public</b> RectangleView(Context context, AttributeSet attrs, <b>int</b> defStyleAttr)
	<b>protected void</b> onDraw(Canvas canvas)
	<b>private void</b> initPaint()
	<b>private void</b> drawRect_2(Canvas canvas)
	<b>public boolean</b> onTouchEvent(MotionEvent event)

#### IV.11 **public class** Setting **extends** LocalizationActivity

Tên biến	final int LANGUAGE_SETTING_RESULT
	Button btnLanguageSetting
	Button btnSoftInfo
	ImageView imvSoftInfo
	boolean imvSoftInfoVisible
	Switch swTheme
	SeekBar seekBarBrightness
	int currentBrightness
	Window window
	ContentResolver contentResolver
Tên hàm	<b>public void</b> onActivityResult( <b>int</b> requestCode, <b>int</b> resultCode, Intent data)

#### IV.12 **public class** SettingsActivity **extends** AppCompatActivity

Tên hàm	private static Preference.OnPreferenceChangeListener sBindPreferenceSummaryToValueListener()
	private static boolean isXLargeTablet(Context context)
	private static void bindPreferenceSummaryToValue(Preference preference)
	protected void onCreate(Bundle savedInstanceState)
	private void setupActionBar()
	public boolean onOptionsItemSelected(int featureId, MenuItem item)
	public boolean onIsMultiPane()
	private static Preference.OnPreferenceChangeListener sBindPreferenceSummaryToValueListener()
	private static boolean isXLargeTablet(Context context)
	public void onBuildHeaders(List<Header> target)
	protected boolean isValidFragment(String fragmentName)

#### IV.13 **public class** Translate **extends** AppCompatActivity

Tên biến	TextToSpeech textToSpeech
	String <b>word</b>
	String definition
	TextView txvWord
	Button btnAddToFavorite
	Button btnDeleteFavorite
	FloatingActionButton <b>btnfabFavorite</b>
	String[] listDefinition
Tên hàm	<b>protected void</b> onCreate(Bundle savedInstanceState)
	<b>public</b> Intent getParentActivityIntent()

#### IV.14 **public class** TranslateText **extends** AppCompatActivity

Tên biến	Button <b>btnStartTranslateText</b>
	EditText edtText
	TextView txvTranslateText

Tên hàm	<b>protected void</b> onCreate(@Nullable Bundle savedInstanceState)
	<b>public void</b> speech( <b>final</b> String text, <b>final</b> String lang)

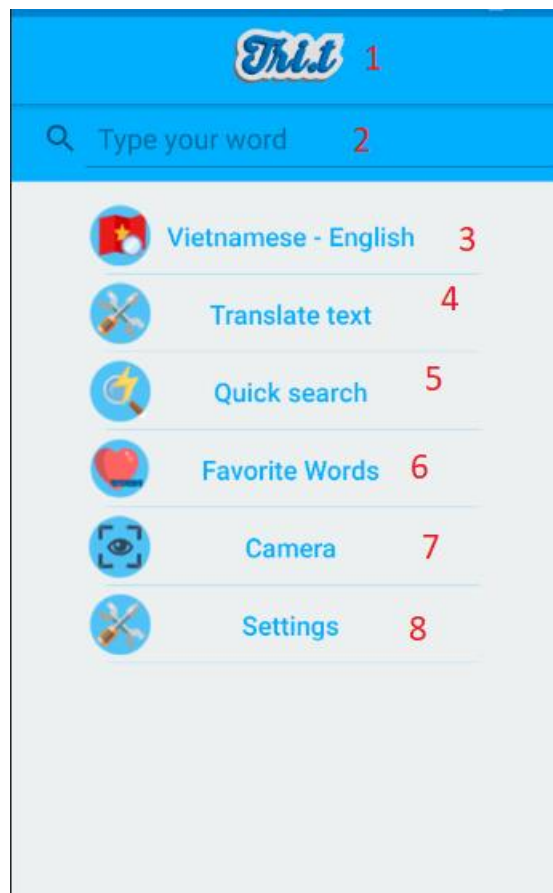
#### IV.15 **public class** VietnameseEnglishActivity **extends** AppCompatActivity

Tên biến	ArrayList<String> <b>arrayList</b>
	ListViewAdapter adapter
	ListView lsv_V_E_Words
	String WORD
	String DEFINITION
	TextView txv_V_E
	FrameLayout frameLayout
Tên hàm	<b>protected void</b> onCreate(@Nullable Bundle savedInstanceState)

#### IV.16 GIAO DIỆN

##### 1. Menu chính

Ý nghĩa : Các menu chính để người dùng truy cập vào các màn hình giao diện



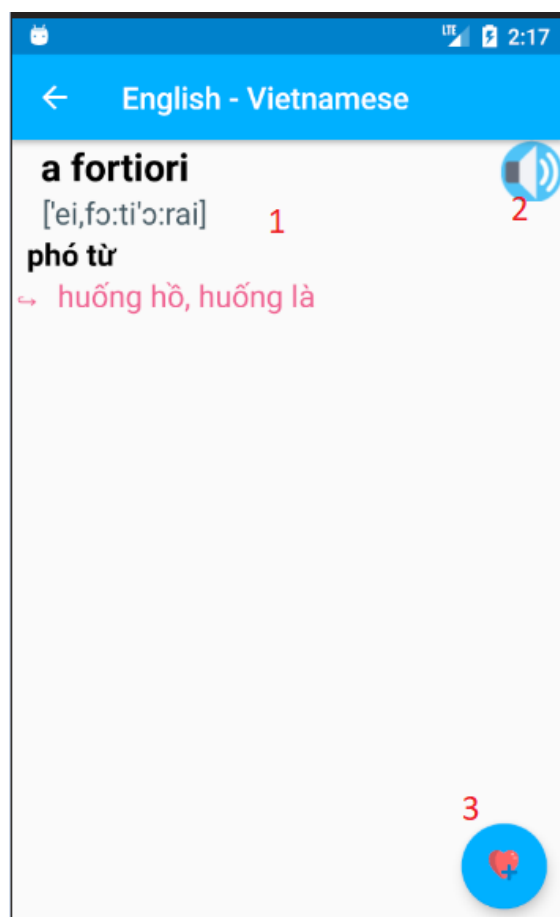
khác

Danh sách các xử lý:



STT	Tên xử lý	Điều kiện gọi thực hiện	Ý nghĩa	Ghi chú
1	Logo	Không thực hiện	Giao diện	Chỉ xem
2	Từ điển Anh Việt	Khi người dùng bấm vào và nhập phím	Vào giao diện màn hình kết quả từ điển Anh Việt	Chỉ bấm
3	Từ điển Việt Anh	Khi người dùng bấm vào	Vào giao diện màn hình tra từ điển Việt Anh	Chỉ bấm
4	Dịch văn bản	Khi người dùng bấm vào	Vào giao diện màn hình dịch văn bản	Chỉ bấm
5	Tra nhanh	Khi người dùng bấm vào	Kích hoạt chế độ tra nhanh	Chỉ bấm
6	Từ yêu thích	Khi người dùng bấm vào	Vào giao diện màn hình từ yêu thích	Chỉ bấm
7	Camera	Khi người dùng bấm vào	Vào giao diện màn hình OCR	Chỉ bấm
8	Cài đặt	Khi người dùng bấm vào	Vào giao diện màn hình cài đặt	Chỉ bấm

## 2. Màn hình kết quả tra từ Anh Việt

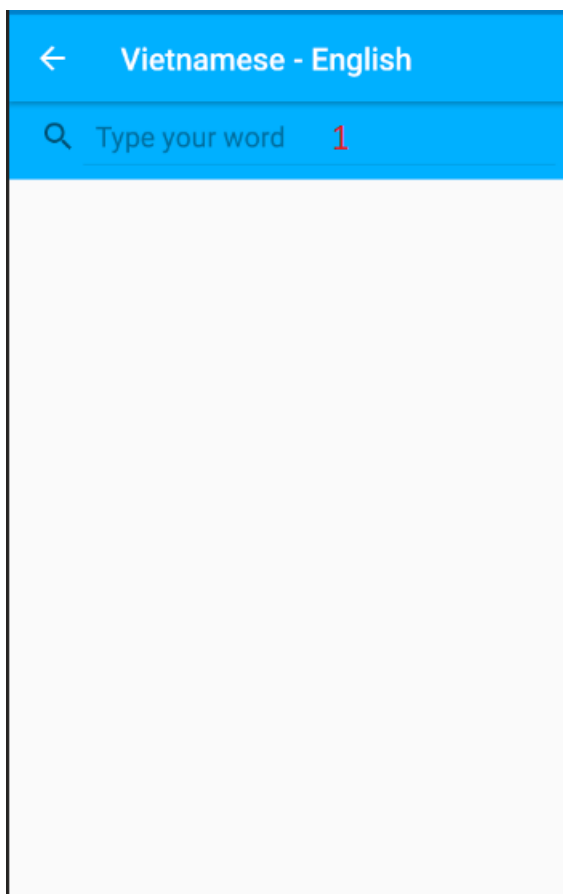


Ý nghĩa : Kết quả tra từ

Danh sách các xử lý:

STT	Tên xử lý	Điều kiện gọi thực hiện	Ý nghĩa	Ghi chú
1	Kết quả	Sau khi người dùng nhập từ và enter	Hiển thị kết quả	Chỉ xem
2	Phát âm	Khi người dùng bấm vào	Phát âm tiếng anh	Chỉ bấm
3	Thêm/xóa từ yêu thích	Khi người dùng bấm vào	Thêm/xóa từ yêu thích	Chỉ bấm

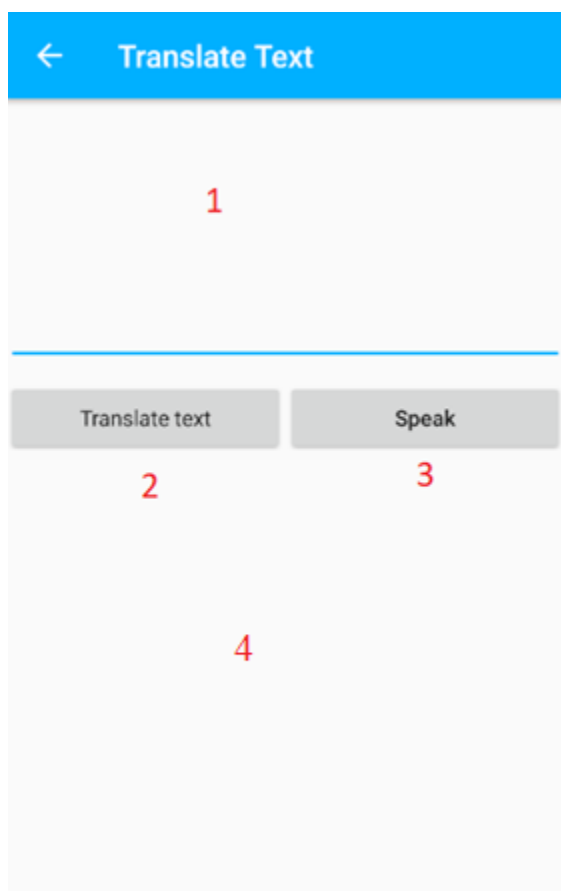
### 3. Màn hình Việt Anh



Màn hình khá giống với màn hình Anh – Việt nhưng ngược lại.

#### 4. Dịch văn bản

Ý nghĩa : Tra văn bản



Danh sách các xử lý:

STT	Tên xử lý	Điều kiện gọi thực hiện	Ý nghĩa	Ghi chú
1	Nhập đoạn văn bản	Khi người dùng nhập từ	Hiện thị đầu vào	Bấm và xem
2	Dịch	Khi người dùng bấm vào	Tiến hành dịch	Chỉ bấm
3	Đọc	Khi người dùng bấm vào	Tiến hành đọc	Chỉ bấm
4	Kết quả	Sau khi người dùng bấm nút dịch	Hiện thị kết quả	Chỉ xem

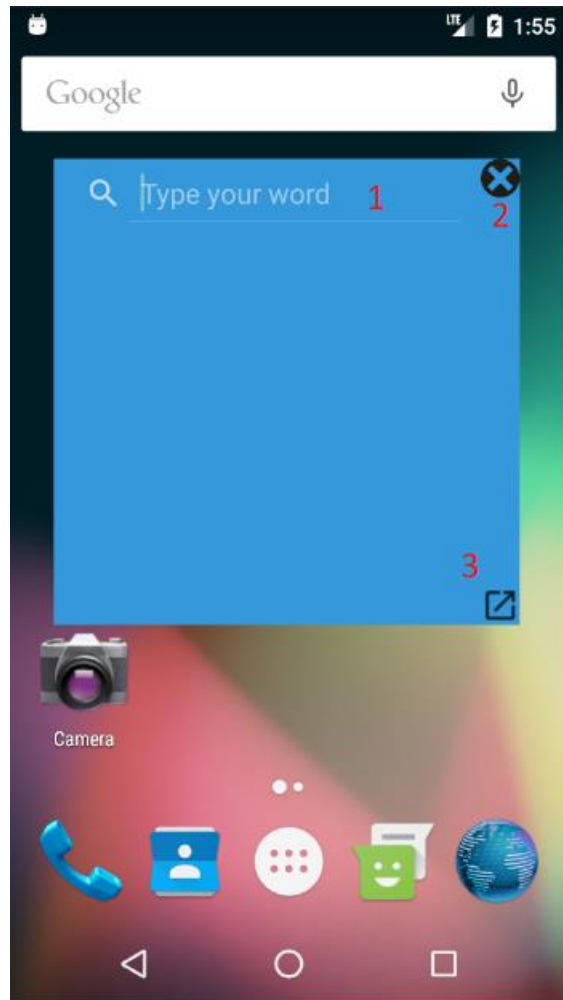
## 5. Tra nhanh

- Thu gọn màn hình sau khi chuyển sang chế độ tra nhanh



Danh sách các xử lý:

STT	Tên xử lý	Điều kiện gọi thực hiện	Ý nghĩa	Ghi chú
1	Icon tra nhanh	Khi người dùng bấm vào	Hiện thị mở rộng của chế độ tra nhanh	Bấm và xem

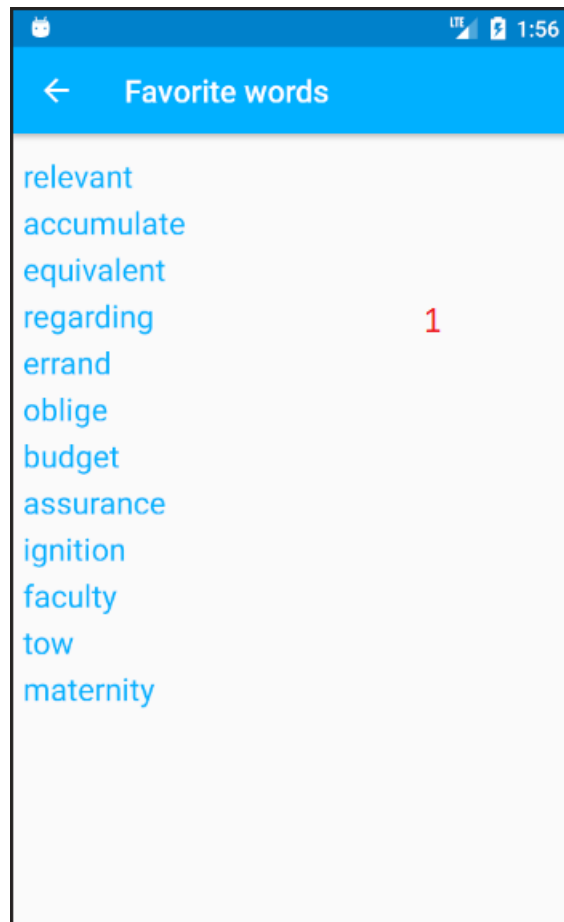


- Mở rộng màn hình tra nhanh

Danh sách các xử lý:

STT	Tên xử lý	Điều kiện gọi thực hiện	Ý nghĩa	Ghi chú
1	Nhập từ	Khi người dùng nhập từ	Hiện thị đầu vào	Bấm và xem
2	Tắt	Khi người dùng bấm vào	Chuyển sang chế độ thu gọn	Chỉ bấm
3	Phóng to	Khi người dùng bấm vào	Chuyển sang chế độ full màn hình	Chỉ bấm

## 6. Từ yêu thích



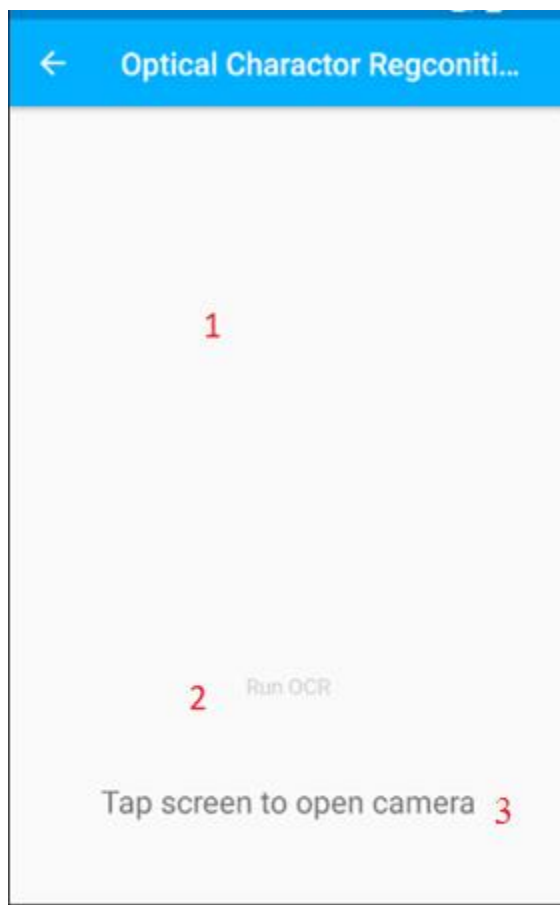
Ý nghĩa: Hiển thị danh sách từ yêu thích

Danh sách các xử lý:

STT	Tên xử lý	Điều kiện gọi thực hiện	Ý nghĩa	Ghi chú
1	Từ yêu thích	Khi người bấm vào	Xem thông tin từ đó	Bấm và xem

## 7. Camera OCR

Ý nghĩa : Kích hoạt chức năng camera OCR



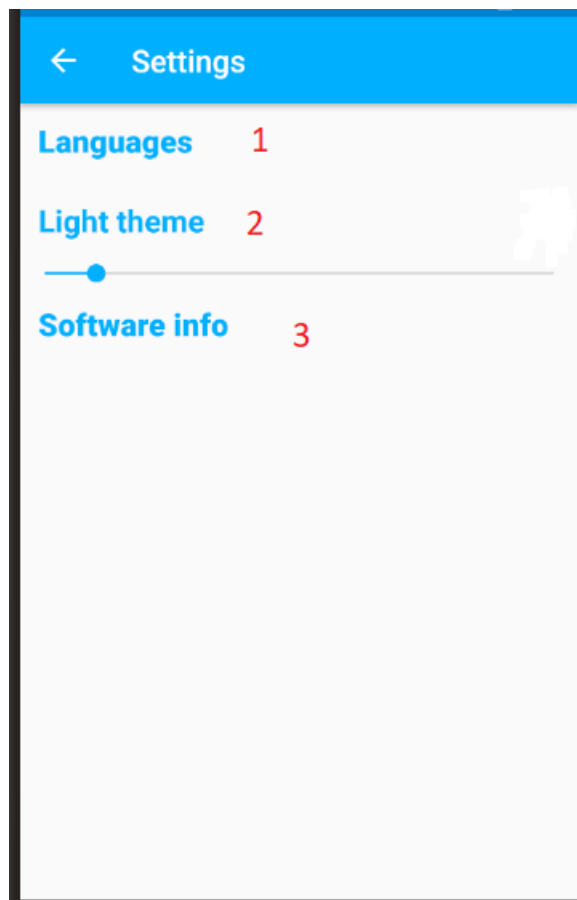
Danh sách các xử lý:

STT	Tên xử lý	Điều kiện gọi thực hiện	Ý nghĩa	Ghi chú
1	Vùng hiển thị	Không thực hiện	Hiển thị hình ảnh	Xem
2	Chạy OCR	Khi người dùng bấm vào	Tiến hành dịch	Bấm và xem
3	Mở Camera	Khi người dùng bấm vào	Chụp hình để dịch	Bấm



## 8. Thiết lập

Ý nghĩa: Thiết lập cài đặt



Danh sách các xử lý:

STT	Tên xử lý	Điều kiện gọi thực hiện	Ý nghĩa	Ghi chú
1	Vùng hiển thị	Không thực hiện	Hiển thị hình ảnh	Xem
2	Chạy OCR	Khi người dùng bấm vào	Tiến hành dịch	Bấm và xem
3	Mở Camera	Khi người dùng bấm vào	Chụp hình để dịch	Bấm

## CHƯƠNG V : CÀI ĐẶT VÀ THỬ NGHIỆM

Môi trường sử dụng Android.

Android là một Hệ điều hành mã nguồn mở và là một hệ điều hành dựa trên Linux cho các thiết bị mobile như Smartphone và máy tính bảng. Ban đầu Android được phát triển bởi Công ty Android với sự hỗ trợ tài chính từ Google, sau đó được Google mua lại vào năm 2005.

Android đưa ra một phương pháp thống nhất để phát triển ứng dụng cho các thiết bị di động, nghĩa là các lập trình viên chỉ cần phát triển Android, và các ứng dụng khác có thể chạy trên các thiết bị khác nhau mà đã được trang bị Android.

Ngôn ngữ cài đặt : Android.

Tên chức năng	Đánh giá hoạt động ( thang 10)	Ghi chú
Dịch Anh Việt	10/10	
Dịch Việt Anh	10/10	
Cửa sổ tra nhanh	8/10	Bị mất focus ở edittext
Dịch văn bản	9/10	Hạn chế vì không tìm được API thích hợp
Sử dụng OCR	9/10	Khó sử dụng, không thành thạo
Sử dụng từ yêu thích	10/10	
Thiết lập	10/10	

## CHƯƠNG VI : KẾT LUẬN

### NHỮNG GÌ ĐẠT ĐƯỢC

1. Cách quản lí một source code lớn, phân chia khối lượng công việc, rèn luyện tinh thần làm việc nhóm, sử dụng thuần thạo dịch vụ cung cấp kho lưu trữ mã nguồn Github,
2. củng cố nền tảng kiến thức kĩ thuật lập trình, lập trình hướng đối tượng, cấu trúc dữ liệu và cơ sở dữ liệu
3. Làm quen với ngôn ngữ Java ví dụ xử lí đa luồng, ...
4. Tiếp cận với nền tảng Android, biết được cách thức hoạt động của hệ thống,
5. Biết cách xây dựng một ứng dụng phần mềm di động đơn giản.

### HẠN CHẾ

1. Chưa xử lí trọn vẹn chức năng dịch văn bản. Vì GOOGLE buộc trả phí API translate. Có suy nghĩ phương án khác sử dụng URL nhưng GOOGLE lại hạn chế việc trích xuất từ file HTML.
2. Chưa xử lí tốt thư viện OCR.

### HƯỚNG MỞ RỘNG

1. Xử lí các hạn chế trong ứng dụng
2. Thêm các trò chơi để giúp người dùng nhớ từ lâu hơn

