ImageTrans Documentation

Xu Lihang

Jun 19, 2021
## Contents

1 Introduction 1

2 Getting Started 3
   2.1 Installation .................................................. 3
   2.2 Verification .................................................. 4
   2.3 New project ................................................. 5
   2.4 Transcription ................................................. 6
   2.5 Translation .................................................. 8
   2.6 Generate the Translated Picture ........................... 11
   2.7 Style Text .................................................. 12
   2.8 Batch .......................................................... 14

3 Text Area Detection and Translation Reinjection 15
   3.1 Text Area Detection ......................................... 15
   3.2 Translation Reinjection ..................................... 17

4 Manga OCR 19

5 Interaction with Third-Party Software 21
   5.1 PSD-related ................................................. 21
   5.2 Others ......................................................... 22

6 Translation Assist 23
   6.1 Corpus ....................................................... 23
   6.2 Sound Words ................................................. 23

7 Tools 25
   7.1 Screen Reader ............................................... 25
   7.2 Silent Translator ............................................ 25
   7.3 Server ......................................................... 26

8 Settings 27
   8.1 General settings ............................................. 27
   8.2 API Settings .................................................. 28
   8.3 Appearance Settings ....................................... 29
Introduction

ImageTrans is an image transcription and translation tool. Unlike Google’s automatic image translation, it allows users to finely control the translation process.

Features:

• Text areas detection (automatic and manual). Supports multiple OCR engines.
• Sort text areas and export the text as xlsx, docx, and other format files
• Use machine translation to pre-translate source text
• Call Photoshop to read PSD files and generate editable translated PSD files
• Text Styling
• Text area detection and text erasing designed for comic translation
CHAPTER 2

Getting Started

2.1 Installation

2.1.1 Full Version

For Windows, unzip the full version to any folder and run ImageTrans.exe. For Mac, drag the app in dmg file to Application.

2.1.2 Cross Platform Version

Download ImageTrans’s zip file, unzip it to any folder, double-click ImageTrans.jar or enter the command line `java -jar ImageTrans.jar` to run it.

The software depends on JRE 1.8. Please download and install it first. Download: 1. Liberica 8u275 full version 2. Baidu Netdisk (code: mhsy)

ImageTrans also relies on OpenCV, please download the runtime file according to your system, unzip and put them under ImageTrans’s folder. Download: 1. GitHub, 2. Baidu Netdisk

2.1.3 Configuration of OCR and Machine Translation

ImageTrans has integrated common online OCR and machine translation APIs. Usually, you need to apply for their API keys to call them. ImageTrans has included the API keys of some services, which can be used directly:

OCR: Baidu, OCRSPACE, Azure

Machine translation: Baidu, Niutrans, Tencent, Google without api key

In addition, some offline OCR and machine translation are also supported.

Offline OCR:

1. Tesseract
If you need to use Tesseract for OCR, please install it by yourself ([Link](#)) and specify the path of Tesseract in ImageTrans. Here is a Windows version for download: [Baidu Netdisk (code: ktpt)](#). Put the downloaded tesseract-ocr with ImageTrans. Put additional language data files under `tesseract-ocr\tessdata`.

2. The built-in OCR in Windows 10

Windows 10 comes with built-in OCR, but you need to install the language environment first. Actually, it has the same engine as Azure and OCRSPACE. This OCR engine in ImageTrans is called WinRT because it is based on the Windows Runtime API.

3. PaddleOCR, EasyOCR

You need to install Python and the above software and run them with the server scripts provided, like the server of PaddleOCR.

4. ABBYY

It supports calling ABBYY Finereader for OCR. You need to specify the path of FineCMD.exe in the Preferences.

Offline machine translation:

1. OPUS-CAT. Opus-CAT is the offline machine translation engine by Helsinki Natural Language Processing Group. Download and install it from here and install the Fiskmo machine translation plug-in to use it.
2. eztrans xp. This is a Japanese-Korean translation software. You can learn about its usage in this issue.

### 2.2 Verification

Each time you run ImageTrans, a validator will show up. You need to verify with your the email and order number. The order number can be found in the email you received after your purchase.
Each email can be used for verification on at most three devices. If you want to use it on the fourth device, you have to reset with your email.

2.3 New project

Click File->New, select a place and enter the project name to save the project file.
2.3.1 Import Pictures

Click File->Import pictures (from a folder) and select in which folder the picture exists to import. It will read all the subfolders and import existing jpg, png files.

You can also use the context menu to paste a image.

It can also import PDF files and extract selectable text.

2.4 Transcription

You can mark text areas in boxes and get their text. The tool supports manual selection and four automatic selection. Further adjustment is also supported.

2.4.1 Create Text Area Boxes Manually

Double-click on the picture to create a selection box. Press on the middle area to move the box and press on the lower right corner to resize.
There is also a quick creation mode. Click the button on the left and you can press and drag your mouse cursor to create boxes quickly.

2.4.2 OCR

Select the text area, select the right language, choose an OCR engine and press OCR.

2.4.3 Detect Text Areas and Recognize Text

Select the language and an OCR engine, click Edit->Detect text areas and recognize text. Youdou detects in paragraph level and other engines detect text lines. You can use the merge buttons at the right side to merge these areas.

2.4.4 Detect Balloons

Select the language and OCR engine (only Baidu and Azure support this), click Edit->Detect Balloons.

Heuristic and natural scene text detection methods are also provided, which are more complex. See Text Area Detection

2.4.5 OCR All Text Areas

We can mark out text areas first and then OCR them in bulk. Click Edit->OCR All Text Areas to do this.

2.4.6 Sort

Sort text areas based on coordinate.

2.4.7 Export

There are several options for exporting.
• Tab-splitted TXT documents include coordinate, font style and text.
• The XLSX file has the same content as TXT’s.
• XLSX file - create worksheets based on folders. It will store picture names, source texts, and target text in separate sheets.
• All text. Store text in separate txt files for each picture.
• Document for translation. Export source text and target text to a docx, txt or XLIFF file.

2.5 Translation

Enter the target text in the translation area and click Save to complete the translation of a text area.
Translations can be exported as docx documents for external translation, and then import back through File->Import translation.
BasicCAT, a computer-aided translation software, supports direct operation of ImageTrans’s project files for translation.

### 2.5.1 Translation memory, machine translation and terminology management

Switch the tab page on the right to Translation Assistant page to use translation memory, machine translation, and terminology management. You need to set up APIs and enable them to use machine translation. You also need to set the language pair of the project, which can be done through Project->Settings->Select language pair.
2.5.2 Pre-translation

Click Project->Batch->Pre-translation to use translation memory or machine translation to pre-translate. Only machine translation is currently supported.
2.5.3 Preview

Click Preview translation in the lower left corner to preview the translation. The precision mode will detect text and reconstruct the background. The imprecision mode will just cover the area with the background color.

When Typesetting mode is checked before preview, target text areas will be boxed out. You can readjust their size and location.

2.6 Generate the Translated Picture

First switch the image scale to 100%, and then click preview to get the translated picture. Click File->Export the current picture to->JPG, and the result will be exported to the out folder in the picture folder. Another option, ORA, supports exporting files to multi-layer image format ORA, which holds layer information for image editing software such as PS, GIMP, and Krita.

In addition, ImageTrans supports exporting as PSDs.
2.7 Style Text

Setting text styles has effects in two processes: Preview translation in ImageTrans and export as PSDs.

Click Project->Settings->Font Style to set the font name, font size, leading, alignment, and so on.

If you want to modify a style, click the style to load the settings and then click Add when the setting is done. Remember to delete the original style. The first style will be the default style.
Because Photoshop requires a special font name, it needs to be obtained from PS. The way is to create a new picture in PS, create a text box, set the font you want, and complete the text editing operation, letting the text box in non-editing status. Then click Read in ImageTrans. Non-Windows systems have to use readFont.jsx scripts.

You can set special font styles for text areas.

Local styles are also supported. Stroke and rotation angles are supported in addition to what the global text style has. When setting the local font style, the interface for the global font style setting will appear. The last new-added style will be saved as the local font style.
Click the Font button on the left to enable the font settings toolbar, which makes it easy to set local styles.

Click the Selection button on the left to enable the multi-selection toolbar. You can adjust the position of multiple text boxes and unify their font styles.

2.8 Batch

All of the above operations on a single picture can be performed in batch through Project->Batch.
CHAPTER 3

Text Area Detection and Translation Reinjection

ImageTrans implements a set of text area detection and translation reinjection methods.

3.1 Text Area Detection

ImageTrans supports four text detection methods: text detection provided by OCR, balloons (bubbles) detection using deep learning object detection, rule-based heuristic detection method and natural scene text detection method.

Here, we mainly explain the heuristic and natural scene text detection methods.

3.1.1 Heuristic

The heuristic method can generate text areas more accurately and provide detailed parameter settings, which can be adjusted for different comics to achieve better results.

Operation:

1. Click Edit->Detect text areas (heuristic) to get all the candidate text areas

2. Click Edit->Text area operation->Get text area confidence. The box color of areas which are unlike text areas will turn to yellow. This kind of area can be removed or hidden. But because some areas are mistakenly identified as non-text areas, it is not recommended to remove them directly.

3. Select the text area to perform OCR and translation. After that, you can remove areas without source text or target text through Edit->Text area operations.
Operations such as OCR will automatically skip low-confidence areas. Because different comics have various sizes, different text area detection parameters need to be set, which can be done through Project->Settings->Text Area Detection. The algorithm details of text area detection can be found here: Rules-based comic text detection method. The confidence of text areas is obtained using a CNN model retrained using TensorFlow’s script. See the code here: https://github.com/xulihang/text-image-classifier.

The operations for text area detection can be done manually in steps via Edit-Text area operations and the merge buttons in the right editing area.

### 3.1.2 Natural Scene Text Detection

The natural scene text detection allows users to call open source natural scene text detection methods such as EAST and CRAFT, which have high accuracy and can detect tilted text. But it generally take a long time to produce the result.

### 3.2 Translation Reinjection

Translation reinjection consists of two steps: source text removal and replacement of target text.

#### 3.2.1 Source Text Removal

There are two modes of text removal. One is the precision mode and the other is the imprecision mode. In precision mode, text mask will be generated first, and then restore the background according to the mask. There are two ways to restore the background: one is to use the image inpainting method, and the other is to use the background color to generate a text mask to cover the text. If the mask is generated incorrectly, it can be modified through Edit->Generate/Edit mask. The mask image will be saved in the picture folder with a mask suffix, while the text-removed image has a text-removed suffix.

**Note:**

1. If the mask is not generated manually, mask and text-removed images are automatically generated every time you switch to the translated version. The generation will takes time.

2. For areas with light font color, you need to set the background color and text color first or enable relevant project settings so that the program will reverse the image color to generate mask correctly.

In imprecision mode, a rectangular box will be generated with a colored background to cover text. If the box has no source text nor target text, it will be transparent. This mode is more suitable for digital images with a simple background.

The following image shows the Mask Editor and the Text Remover, which can be used to adjust the mask and text removal results:
There is plug-in support for mask generation and image inpainting to use third-party methods. The existing plug-in is Sickzil-machine.

### 3.2.2 Translation Replacement

You can adjust the location and size of the target text boxes. The font size will be automatically adjusted. You can set this in the project settings.

### 3.2.3 Color Detection

This tool can automatically detect the background color and text color roughly. Click Edit->Color operations to do the detection.
The text of Japanese manga is arranged vertically from right to left, and there will be Furigana for phonetic notation, so the processing methods will be different.

The following is the related features of ImageTrans for manga OCR.

1. Strip Furigana

Check Strip Furigana on the OCR toolbar to remove Furigana in the image during OCR.

Original image:

After:

2. Vertical to Horizontal

Check Vertical to Horizontal in the OCR toolbar so that the text will be rearranged in during OCR. In this way, many OCR engines that can only recognize horizontal Japanese can also be used.

Original image:
3. Parameter Setting of The Heuristic Text Detection Method

The default heuristic text detection method merges horizontally and then vertically. For manga, it can be set to merge vertically and then horizontally. The overlapping ratio of text lines can also be adjusted to avoid the problem of merging different paragraphs.

4. Right to Left Reading Order

Check this option in project settings so that when merging text areas, the text on the right will be placed in the front.
5.1 PSD-related

ImageTrans can interact with Photoshop using scripts.

Note: If you are on a Windows system and have a full version of Photoshop installed, ImageTrans provides exe scripting tools that automatically open PS. In other cases, you need to open the PS yourself and click File->Script->Browse to run the scripts.

Photoshop Script Repository: https://github.com/xulihang/ImageTrans_PhotoshopScripts

5.1.1 PSD to JPG

The software does not support PSD files directly. You need to convert PSDs to JPGs first.

Open the conversion tool via Tools->PSD to JPG. It will call Photoshop to do the bulk conversion.

Non-Windows systems have to use the provided JSX script: pdf2jpg.jsx. Select the directory that needs to be converted.

5.1.2 Read Text Areas from PSDs

If you are working on a PSD file, you can read the coordinates, size, and text of text layers in the PSD file, display it in ImageTrans, and automatically replace the text with translation without generating an overlay layer. Click the Edit->Read text layers from PSD files, if present.

Non-Windows systems have to use the provided JSX script: readTextLayers.jsx. Select the folder to export, and then click Project->Import text areas exported with PS scripts to import.
5.1.3 Read Font Name

The font name is in a PS-specific format, which can be obtained using readFont.jsx. The script reads the font information of the first text box.

5.1.4 Generating PSD

After the translation is completed, you can generate a PSD file for further adjustment in PS. Click File-Generate editable PSD files to bring up the generating options dialog box.

Here’s a description of the options:

- PSD exists - directly process the original PSD file. PSD files have to be put together with the JPG files. If it is not checked, new PSD files will be generated based on the JPG files.
- Replace with translation - Otherwise the source text will be used.
- Add overlay mask - Add overlays to cover the source text. If the text area corresponds to a text layer in the PSD, no overlay will be added.
- Use Precision Mode - The overlay in imprecision mode is a rectangular box. In precision mode, the software will accurately remove text.
- Flip image horizontally - for Chinese Comics Translated into Japanese Comics
- Use point text - All text boxes will use point text.
- Export only - Export data files only. Do not run Photoshop.

Please patiently wait until the prompt window pops up indicating the operation is done. During the operation you can switch to the window of PS to view the operation in progress. If PS displays relevant dialogs, you need to handle them manually.

Non-Windows systems have to use this script: addLayers.jsx.

5.2 Others

Use File->Export to export text as a docx document for others to translate. ImageTrans can reimport the translation.

In addition, the tool’s project files are stored in json format and you can write your own programs to handle them.
ImageTrans supports common functions in computer-assisted translation software, including translation memory, terminology management, machine translation and concordancing. It also has a comics sound words search engine. Here is a detailed introduction to the corpus and sound words search engine.

6.1 Corpus

You need to store multilingual data in TMX format before uploading them to the system. Currently, a corpus created from the subtitles of ten seasons of Friends is provided for learning purpose.

See here for more details: Online Translation Memory Search

6.2 Sound Words

You need to create your own sound words dictionary and store it in a XLSX file, with entry names, segmented words (for Chinese), sound attributes, and explanations before uploading it to the system. The sound word entries of the New Oxford English-Chinese Chinese-English Dictionary are currently available for learning purposes.

See here for more details: Comics Translation - Sound Words
CHAPTER 7

Tools

ImageTrans also comes with a range of tools.

7.1 Screen Reader

Make screen captures, OCR and translate. Results can be saved to the project.

7.2 Silent Translator

The pictures can be translated in batch silently. It can be used as in command line.

How:

```java
java -jar ImageTrans.jar configPath usePrevious detectOnly outdir fileListPath
```
Parameters:

- configPath: the path of the parameter file. If you open Silent Translator you can see the parameters on the right side.
- usePrevious: whether to use the previous data. The optional values are true and false
- detectOnly: whether to detect only text areas. The optional values are true and false
- outdir: output folder
- filelistPath: a list of files to be processed, separated by newlines

It can be used with *ImageTrans_Server* <https://github.com/xulihang/ImageTrans_Server> to provide an online service.

### 7.3 Server

This tool allows other software to call ImageTrans to translate images. *ImageTrans Chrome Extension* relies on this tool.

The extension can use ImageTrans to translate images on webpages. Images and processing results will be automatically added to the current project of ImageTrans.
8.1 General settings

8.1.1 Balloon Detection

The tool’s balloon (bubble) detection uses object detection services provided by Baidu EasyDL and Azure. You need to create your own account and train your own model.

Recommended labeling methods:

Box text areas in all balloons. Let the box fits the text as closely as possible.
After the model is published, you can set the URL through File-Preferences-General setting. The corresponding API key is also required.

It also supports calling the offline bubble detection model trained by Darknet Yolo or TensorFlow Object Detection API.

### 8.1.2 Text Image Recognition

By default, the software uses the local model to recognize text and non text areas. You can also call an API service.

The default address is: http://127.0.0.1:8082/classify

Please run the code in this repository: https://github.com/xulihang/text-image-classifier.

### 8.1.3 Natural Scene Text Detection

EAST and CRAFT natural scene text detection methods are supported.

The default address is: http://127.0.0.1:8080/detect


### 8.2 API Settings

Click File->Preferences->API to set up APIs for OCR and machine translation services provided by Baidu, Youdao, Microsoft, Tencent, etc.
The following is a list of services that need to set the API key.

OCR:

- Google
- Sogou
- Baidu
- Tencent
- Youdao
- OCRSPACE
- Microsoft Azure

For Machine translation, see BasicCAT’s docs.

### 8.3 Appearance Settings

Go to File->Preferences->Theme to adjust the appearance. In addition to the default theme, there are black and green themes.

You can also use external CSS to adjust the appearance of the software.

For example, the following CSS can control the text size of the text editing areas:

---

8.3. Appearance Settings 29