

EXTENSION
WEB OCR

SEPTEMBER 2024





CONTENTS

MODELS

DEMO

FUTURE
UPDATE

INTRODUCTION

IDEA

PROBLEMS

Q&A

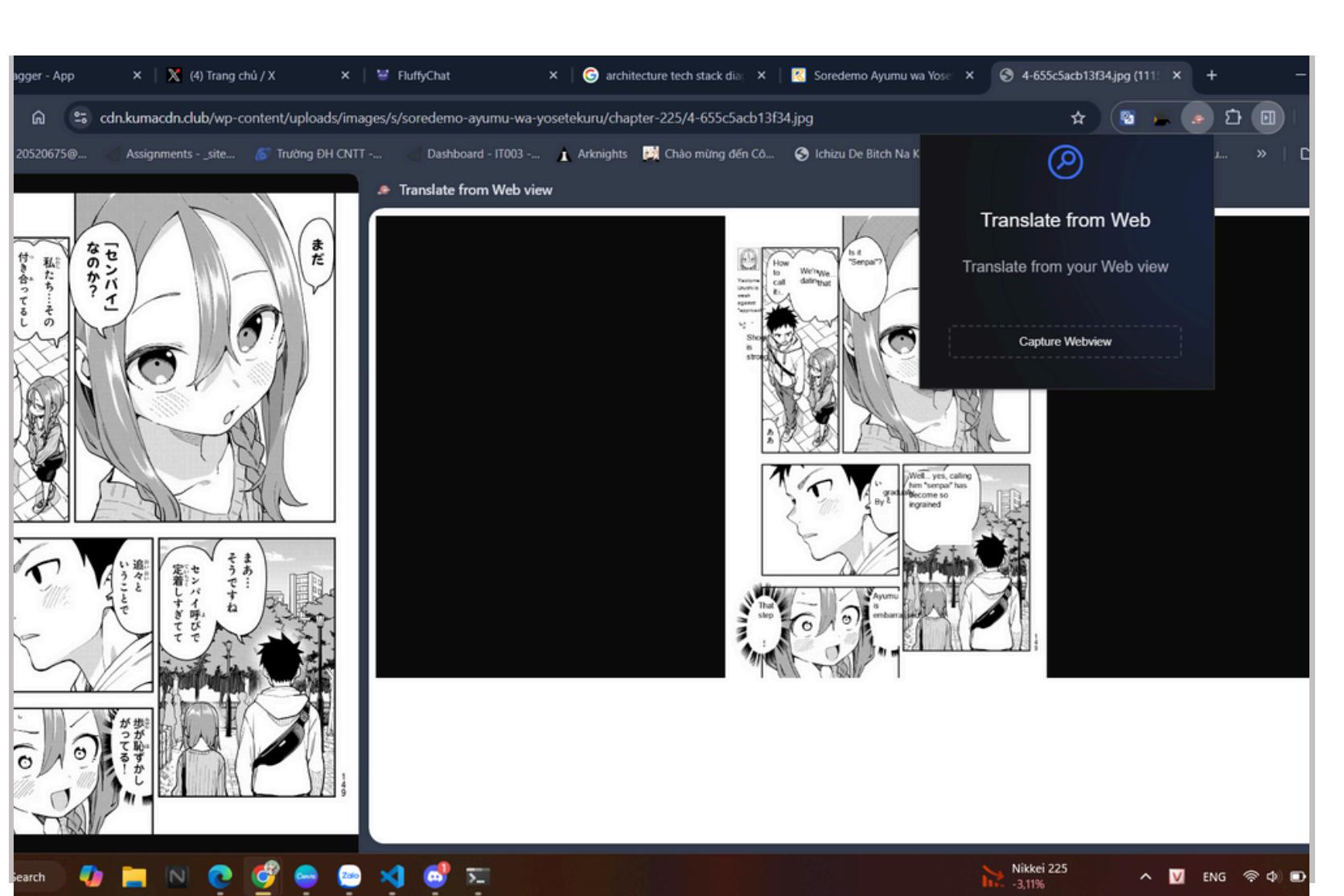




Introduction

This is an extension effortlessly detects and translates text from your web view, saving you time and effort.

- Run on browser
- Almost real-time





MODELS

EASYOCR

EasyOCR is a python module for extracting text from image. It is a general OCR that can read both natural scene text and dense text in document.

MANGAOCR

Optical character recognition for Japanese text, with the main focus being Japanese manga.



EasyOCR

EasyOCR is a versatile Python library for OCR that simplifies the process of extracting text from images. Ideal for automating data entry and image analysis tasks, offers exceptional accuracy and flexibility

- High recognition rates
- Support for multiple languages
- Efficient processing
- Optimized for different hardware



MangaOCR

MangaOCR is a Python library specifically designed for extracting text from Japanese manga images. It's a powerful tool for researchers, developers, and enthusiasts who work with manga content.

- High accuracy
- Specialized for manga
- Robustness to noise and artifacts
- Efficiency



Idea





IDEA



Detect text
with EasyOCR

Extract text
from image
with MangaOCR

Translate
extracted text

Add translated
text to image



IDEA



DETECT TEXT WITH EASYOCR



EasyOCR

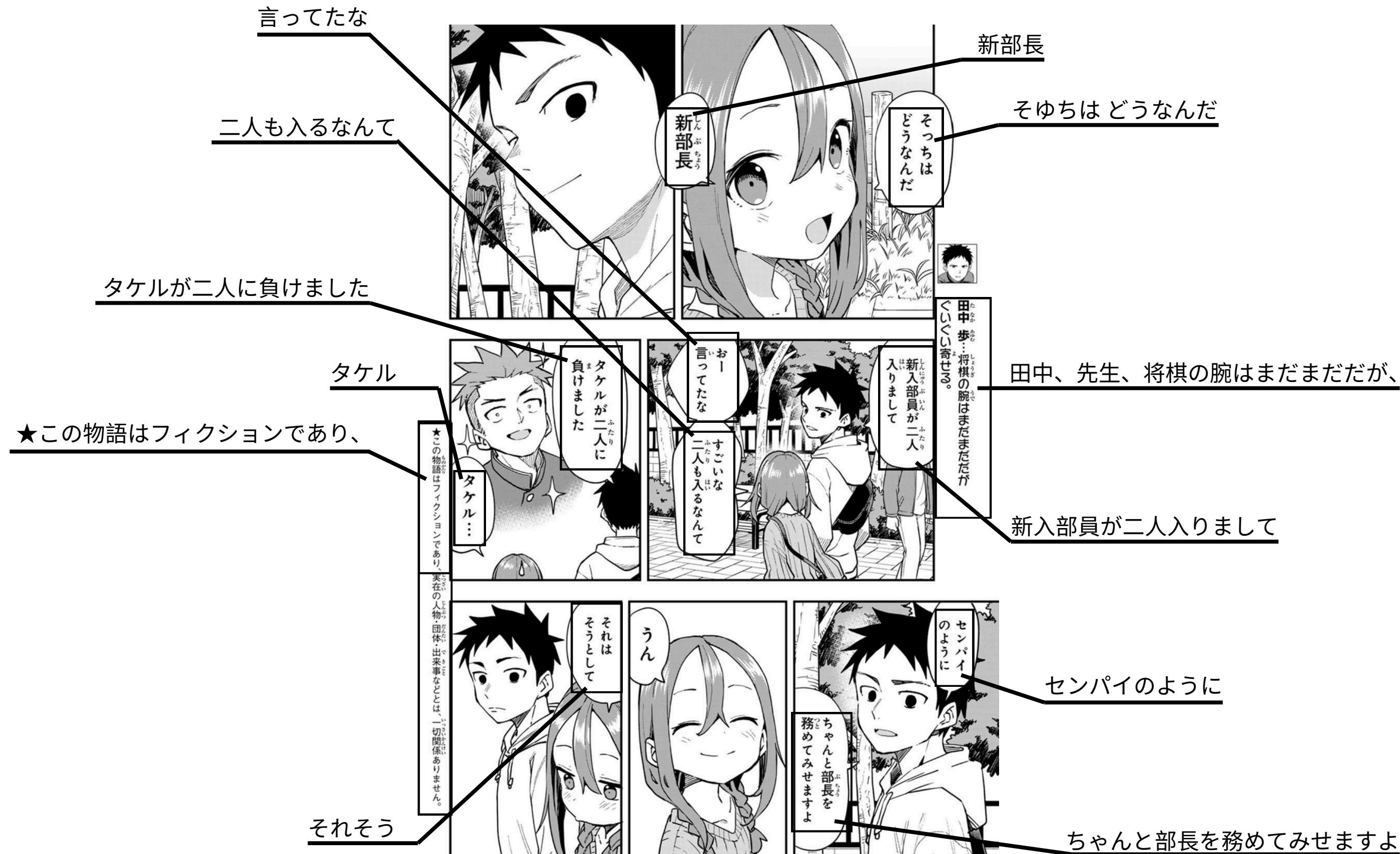




Idea



Extract text from image with MangaOCR





Translate extracted text



それそう---->That's right

新入部員が二人入りまして---->Two new members joined

どうなんだ---->What do you think?

実在---->Real

★この物語はフィクションであり、---->★This story is fiction.

して---->do

切開係---->incision system

新部長---->new minister

ぐいぐい寄せる。---->Pull closer.

田中、先生、将棋の腕はまだまだだが、---->Tanaka, sensei, your shogi skills still need improvement,

タケルが二人に負けました---->Takeru lost to the two of them.

タケル---->Takeru

ちゃんと部長を務めてみせますよ---->I'll show you how to act as a manager.

そゆちは---->Soyuchi is

ません。---->No.

二人も入るなんて---->Two people can enter

センパイのように---->Like Senpai

出来事などとは、---->Incidents and such are

人物・団体---->Individuals/organizations

言ってたな---->You said.

Idea

>>>

Add translated text to image

Translated texts

それそう---->That's right

新入部員が二人入りまして---->Two new members joined

どうなんだ---->What do you think?

実在---->Real

★この物語はフィクションであり、---->★This story is fiction.

して---->do

切開係---->incision system

新部長---->new minister

ぐいぐい寄せる。---->Pull closer.

田中、先生、将棋の腕はまだまだだが、---->Tanaka, sensei, your shogi skills still need improvement,

タケルが二人に負けました---->Takeru lost to the two of them.

タケル---->Takeru

ちゃんと部長を務めてみせますよ---->I'll show you how to act as a manager.

そゆちは---->Soyuchi is

ません。---->No.

二人も入るなんて---->Two people can enter

センパイのように---->Like Senpai

出来事などとは、---->Incidents and such are

人物・団体---->Individuals/organizations

言ってたな---->You said.





DEMO



ENCOUNTERED PROBLEMS

Detect text in image

Combine bounding boxes

Capture Webview

Passing data

Docker Deployment



Detect text in image

	Vertical line	Speed	Detect texts	Fine-tune	Size
PaddleOCR	✓	1	2	1	1
TesseractOCR	✗				
EasyOCR	✓	2	1	2	2



DETECT TEXT IN IMAGE



PaddleOCR

EasyOCR



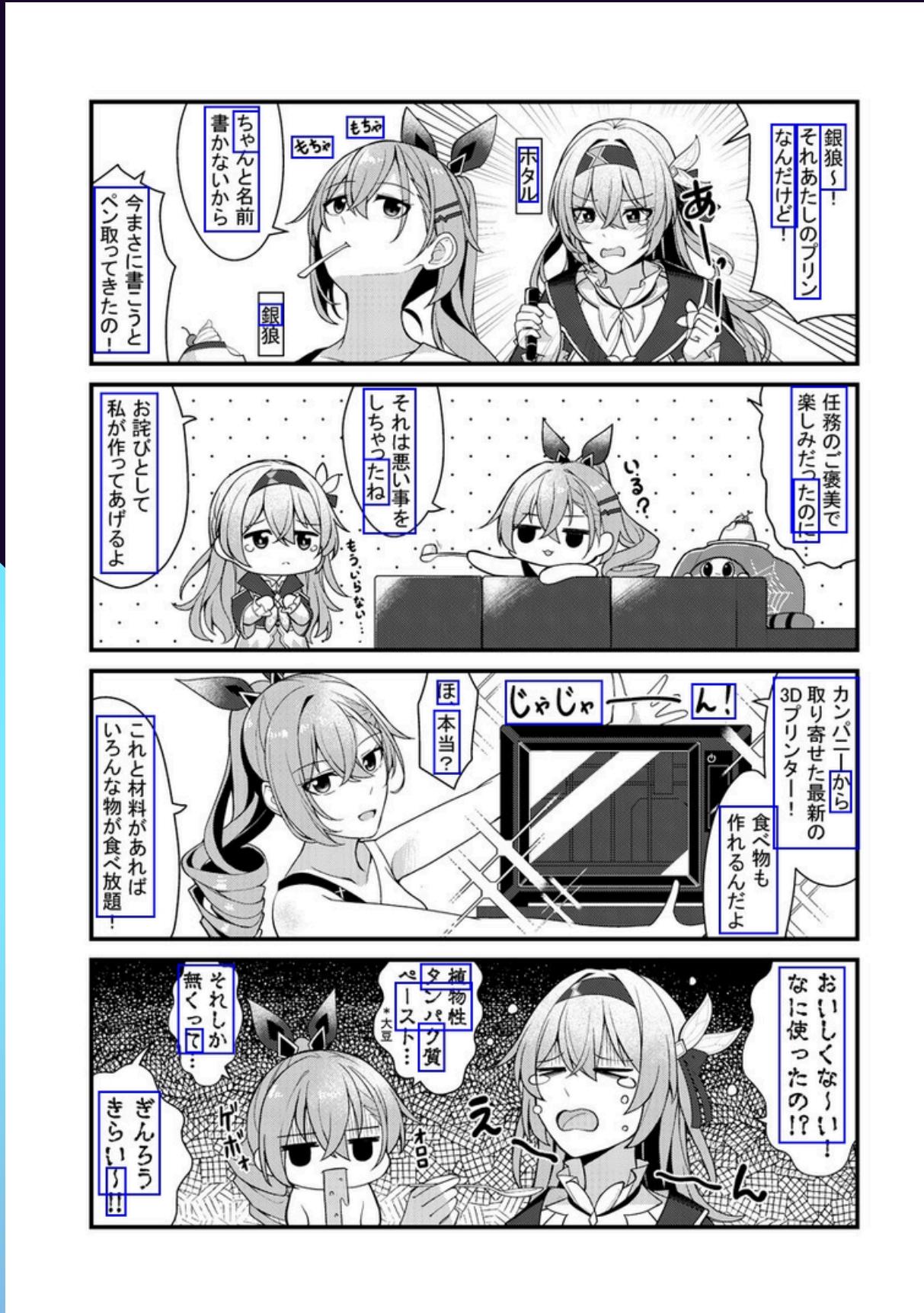
DETECT TEXT IN IMAGE



PaddleOCR



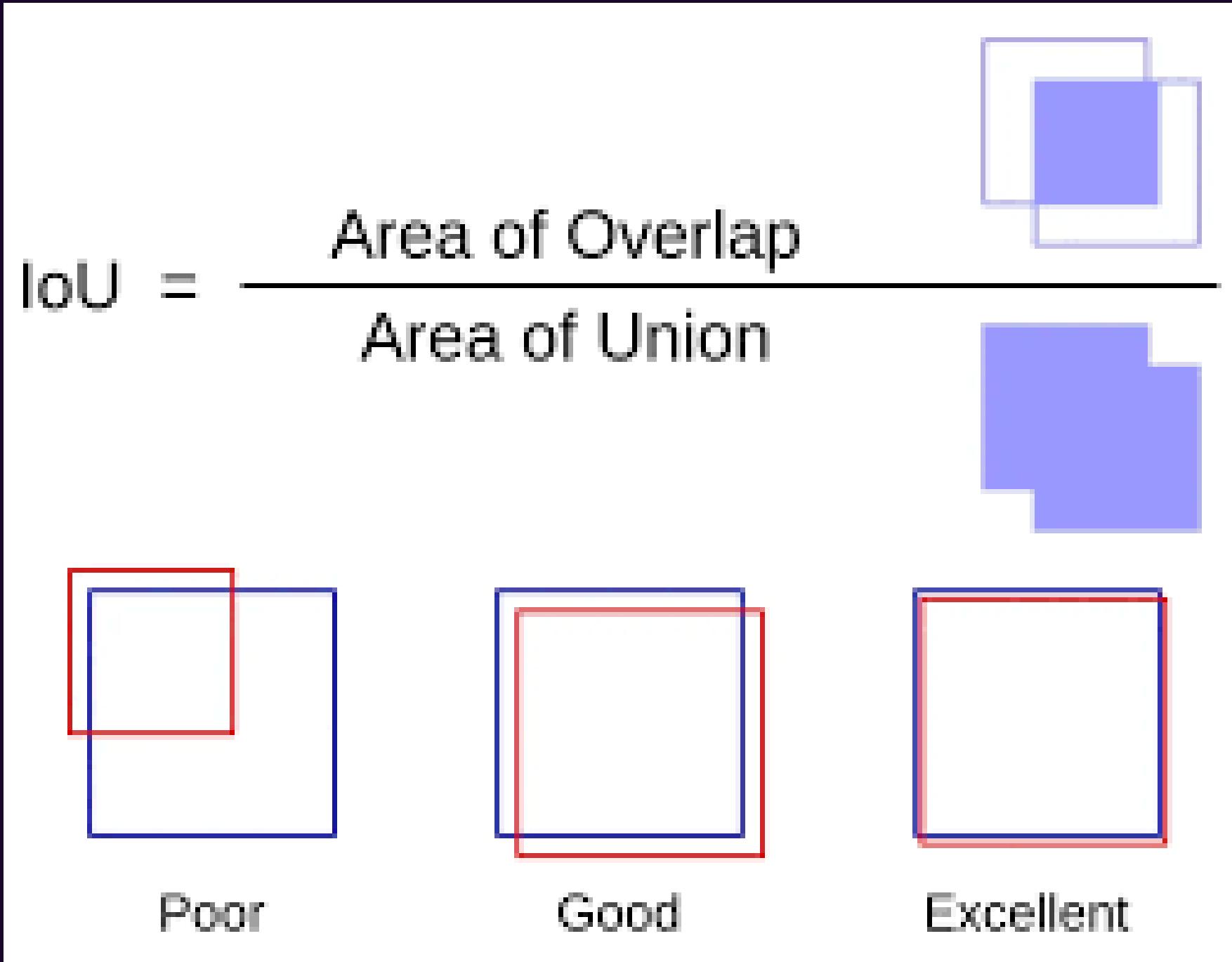
EasyOCR



Combine bounding boxes

COMBINE BOUNDING BOXES

Solution

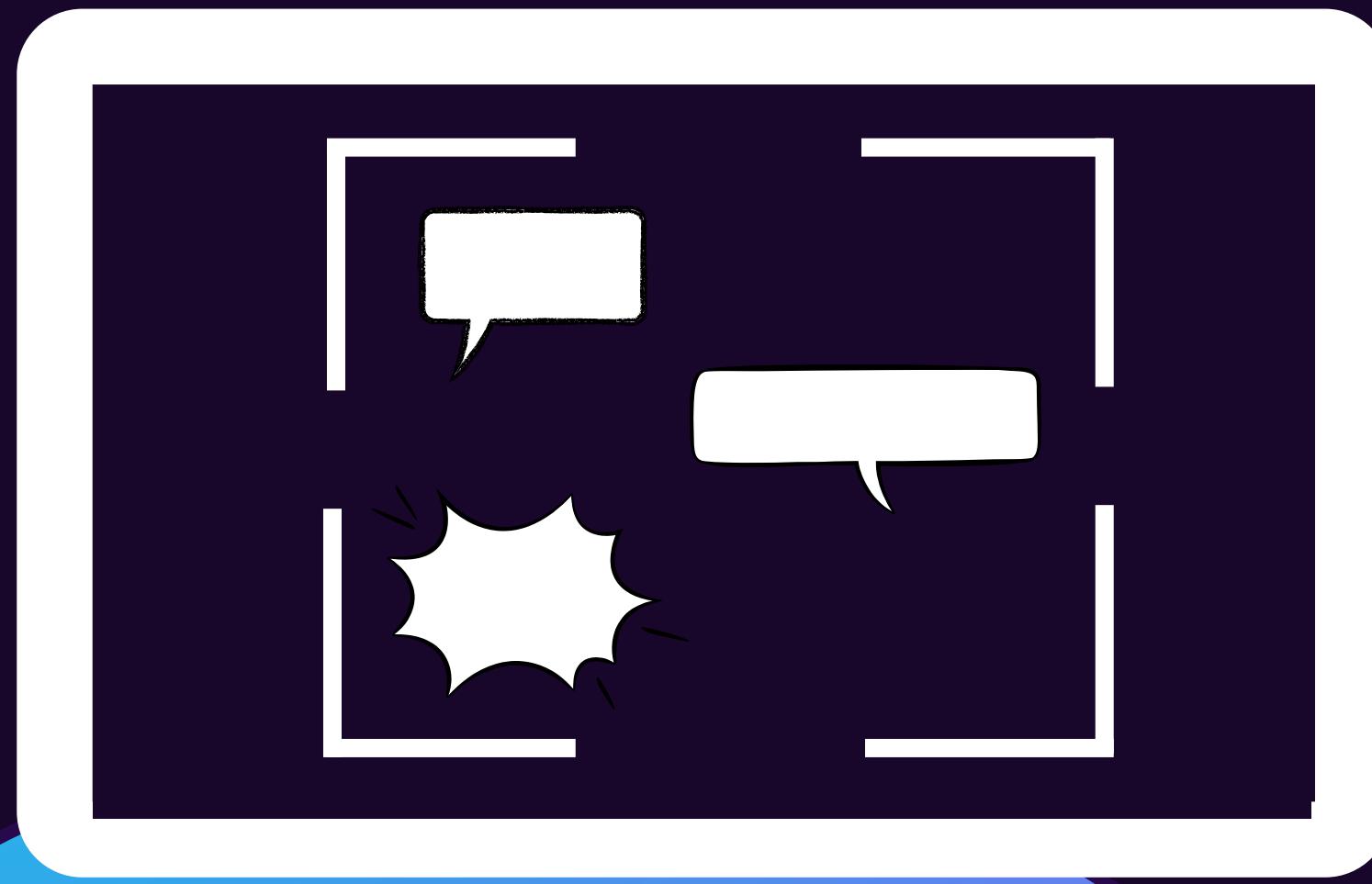


Calculate the IoU of each bounding box with the remaining bounding boxes

If $\text{IoU} > 0$:
Merge 2 bounding boxes



Capture WebView



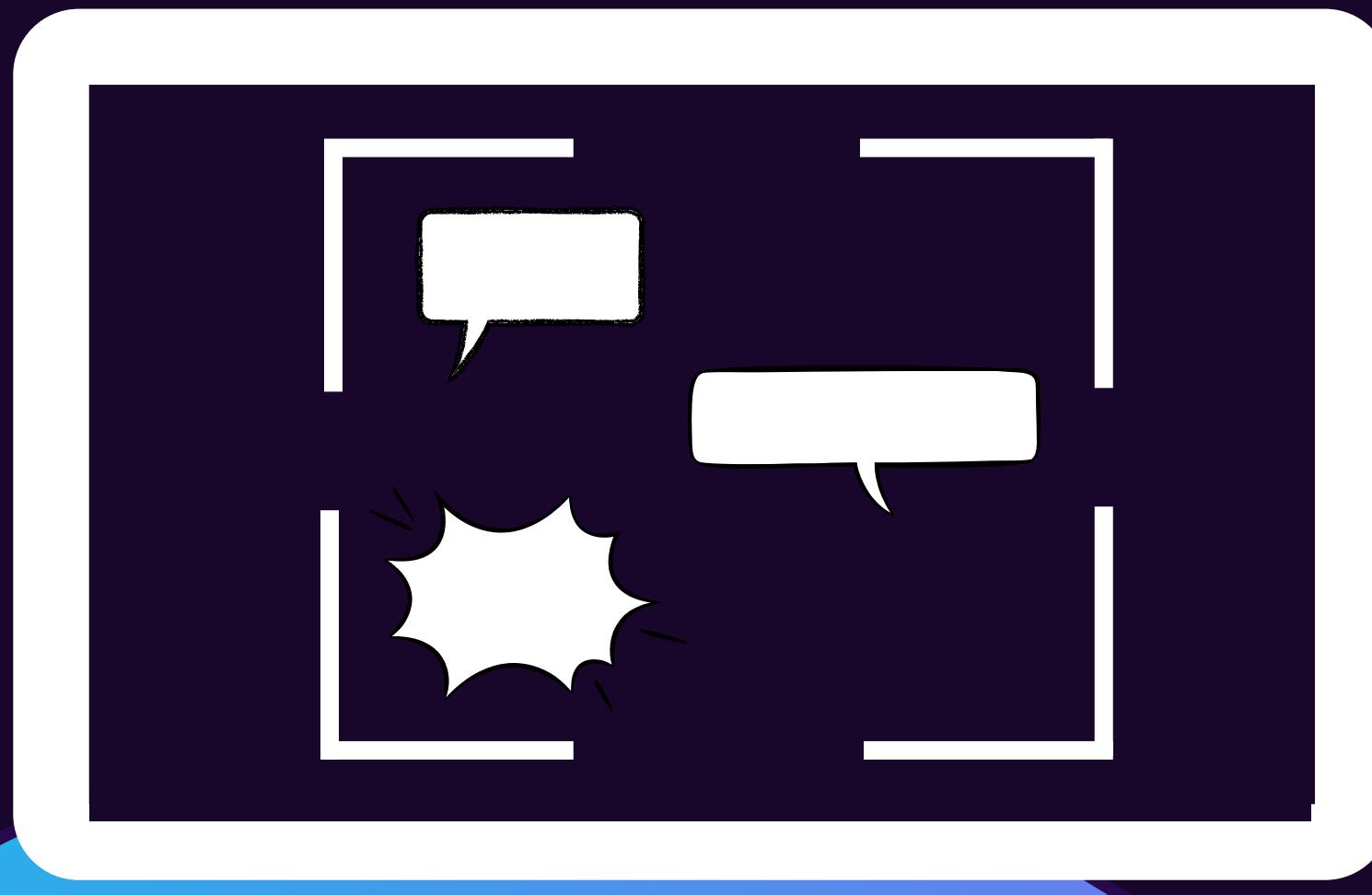
- **html2canvas**

The script allows you to take “screenshots” of webpages or parts of it, directly on the users browser. The screenshot is based on the DOM and as such may not be 100% accurate to the real representation as it does not make an actual screenshot, but builds the screenshot based on the information available on the page.

X ISSUE: Cannot capture image from tag X



Capture WebView



- **Chrome extension API**

An Extensions API consists of a namespace containing methods and properties for doing extensions work, and usually, but not always. Methods in extension APIs are asynchronous unless stated otherwise. Asynchronous methods return immediately, without waiting for the operation that calls them to finish.

✓ Support capture display tab screen ✓



Passing data

Client



data:image/jpeg;base64,/9j/4AAQSkZJRgABAQAAAQABAAAD/4gHYSUNDX1B...YBgGA/9k=



data:image/jpeg; base64,iVBORw0KGgoA...U5ErkJgg==

Server

/9j/4AAQSkZJRgABAQAAAQABAAAD/4gHYSUNDX1B...YBgGA/9k=



iVBORw0KGgoAAAANSUhEUgAAB4AAAANICA...U5ErkJgg==



Docker deployment

Issue: Taking up to 10.25GB for installing required images
(Pytorch image took 7.6GB)

Reason:

- EasyOCR and MangaOCR using torch library to interact with GPU/CPU of computer -> Pull Pytorch image

Solution: Free up memory space (2 ways)

- Delete unnecessary data, applications, etc...
- Using **Diskpart** Command





FUTURE UPDATE

- 01 Translate to other languages
- 02 Using other translate model or API to improve translation quality
- 03 Changed the UI to be more user-friendly
- 04 Build software to run in the background on the computer
- 05 Can detect and recognize languages other than Japanese and Chinese



TRAN QUANG NHAT



Q&A



**THANK
YOU
FOR
LISTENING**