University Of Information Technology, VNU-HCM
CS117.N21.KHTN Computational Thinking

MEDICINE PILL IMAGE RECOGNITION

TEACHER INSTRUCTION:
PhD. Duc Thanh Ngo
thanhnd.@gm.uit.edu.vn

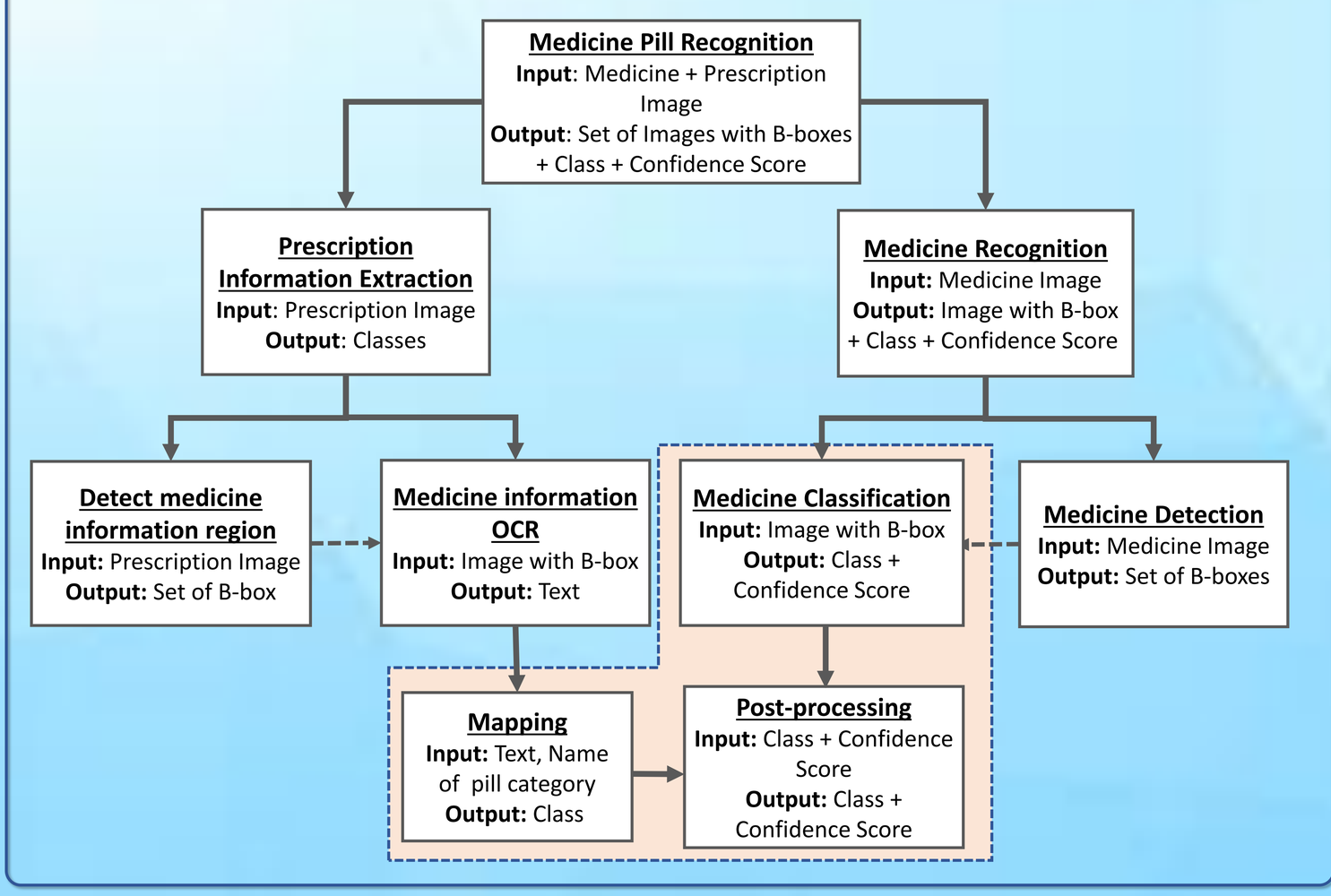
TEAM MEMBERS:
Viet Nhat Nguyen 21520378.@gm.uit.edu.vn
Hoang Thuan Nguyen 21521501.@gm.uit.edu.vn
Nhat Tung Mai 21520523.@gm.uit.edu.vn
Trong Tai Ha 21520436.@gm.uit.edu.vn
Minh Quan Vo 21520093.@gm.uit.edu.vn

Introduction

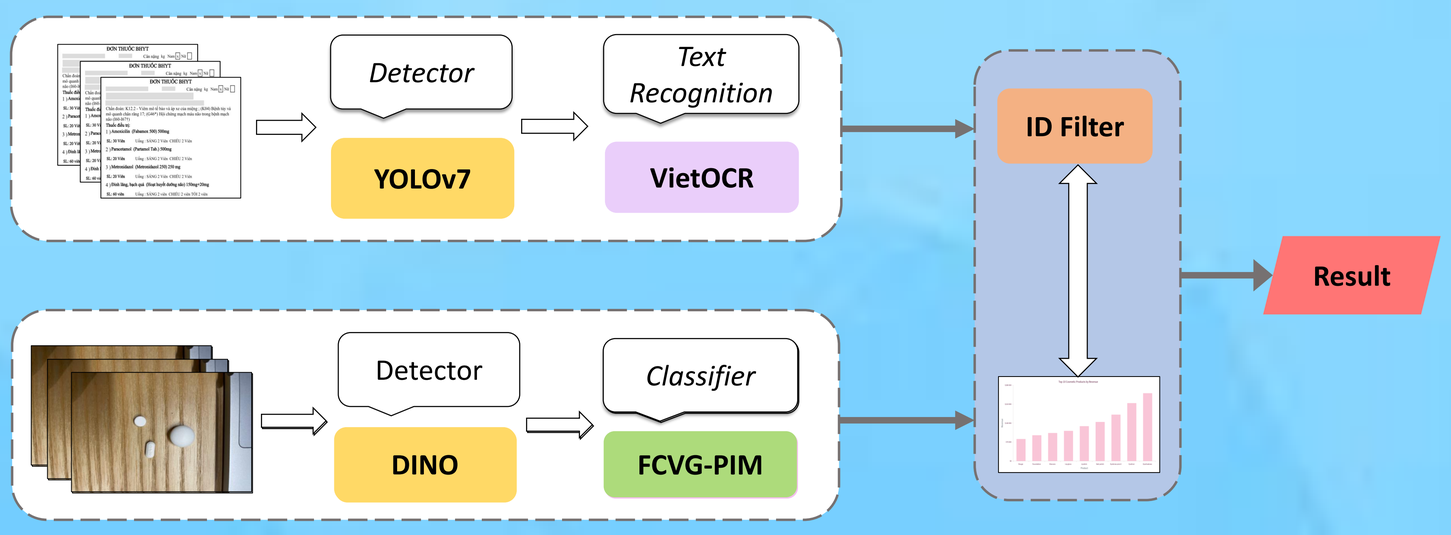
Medications are used to support patients, helping them alleviate or treat their ailments. However, taking medication incorrectly can lead to serious consequences, including reduced treatment effectiveness, adverse effects, or even death. According to the World Health Organization (WHO), one-third of all deaths are caused by medication errors rather than the underlying diseases. The widespread development and increasing demand for medication consumption have escalated the need for accurate identification of drugs with the help of supportive applications.

In this poster, we introduce the model for medicine pill image recognition. It aims to address the aforementioned issue. Consequently, it can help minimize medication errors, enhance effectiveness in treatment process, and facilitate prompt recovery for patients.

Decomposition



Algorithm



Experiment

Models	Options		mAP	m$AP50$	$mAP75$	$mAPs$	$mAPm$	$mAPl$
	#class	top_k						
DINO + FGVC-PIM	108	2	0.67	0.89	0.77	**0.40**	0.63	0.67
DINO + FGVC-PIM	108	3	0.68	0.91	0.79	**0.40**	0.64	0.69
DINO + FGVC-PIM	108	4	**0.69**	**0.92**	**0.80**	**0.40**	**0.66**	**0.70**
DINO + FGVC-PIM	108	5	0.68	0.91	0.80	**0.4**	0.65	0.68

Abstraction / Data Presentation

Abstract: Using a prescription image along with one or more related images of pill associated with that prescription, the model returns bounding boxes indicate locations for each pill in each image as well as its name, with the outside pills named **unknown**.

Input:

- **Prescription image (Printed prescription):**
  - Frontal view, not distorted with no background
  - Background conditions:
    - Sufficient lighting (good distribution histogram of the image)
    - Resolution: $1600 \times 1150, 1420 \times 770, 1070 \times 760$
  - According to the Circular of the Minister of Health (2018/TT-BYT)
  - The containing types of medications are from a predefined list (107 types)
- **Pill images:**
  - Color images, frontal view
  - Each image may contain one or more pills (of the same or different types), occlusion may be occur
  - The pills must be **intact**
  - Wide range of lighting conditions, backgrounds and resolutions

Output: A set of bounding boxes indicates location of each pill in each image as well as the name of each pill, with the outside pills named **unknown**

Pattern Recognition

Object detection:

- **Detect Medicine Information Region**: with fixed ROI, we can use object detection to locate regions containing the pill's name.
- **Medicine Detection**: Define the location of pills in the image
  - Both of the mentioned problems are similar in input and output in object detection problem:

Input: Image
Output: Set of bounding boxes contain objects in image.

Object Classification:

- **Medicine Classification**: Apply solution of classification problem with 108 classes (107-th class is pills not in prescription)

Input: Image with bounding boxes or Cropped Image
Output: Class of object

Optical Character Recognition – OCR:

- **Medicine infomation OCR**: Map information from image to text information that fits the idea of OCR problem.

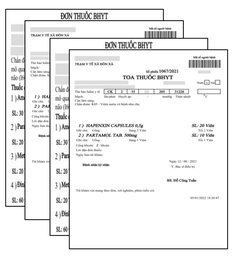
Input: Image with bbox
Output: Text

Furthermore, for other sub-tasks like **mapping**, we can use the method of measuring similarity between two strings based on the **Levenshtein metric**. Or for **post-processing**, we can apply the **ensemble result** technique to combine multiple predictions from different models for better effectiveness.

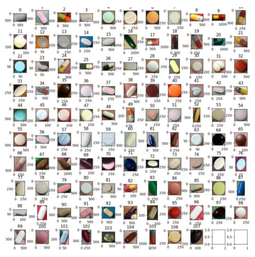
Dataset

Dataset is taken from VAIPE Challenge. Includes **1171 prescription images** and **9500 pill images** belong to 108 classes. The 107-th class contains pill images which are not in prescription.

Prescription images



Pill images



Metrics

IoU

$$\frac{intersection}{union} = \frac{A \cap B}{A \cup B}$$

Determine whether a prediction is considered a **True Positive** (TP) or a **False Positive** (FP)

mAP

$$\frac{1}{|classes|} \sum_{c \in classes} \frac{|TP_c|}{|TP_c| + |FP_c|}$$

Evaluate the **overall performance** of an object detection model.

Conclusion

After applying computational thinking to solve the Medicine Pill Recognition problem, the result of the process is acceptable. In particular, the model can handle small objects, obscured objects, or overlapping objects. Although the object detection model performs quite well, the result achieved from the classifier is still limited. The reason is that there are many types of pills with high similarity. Therefore, to apply this topic to real life, the accuracy needs to be further improved.