INT3404E 20 - Image Processing - Group 3
Sino-nom character localization report

Hoang Duc Anh
Tran Thi Van Anh
Nguyen Thi Thuy Huong
Tran Hanh Uyen

1 Team Leader

Nguyen Thi Thuy Huong

2 Planning and Proposal Development

2.1 Conducting a Literature Review and Survey

We conducted a thorough literature review and survey of existing solutions or methods relevant to the chosen problem. The state-of-the-art methods can be categorized into two main types: one-stage methods and two stage-methods:

1. One-stage methods prioritize inference speed, and example models include YOLO, SSD and RetinaNet.
2. Two-stage methods prioritize detection accuracy, and example models include Faster R-CNN, Mask R-CNN and Cascade R-CNN.

We surveyed some state-of-the-art (SOTA) Object Detection Models at roboflow.com and paperswithcode.com. Among numerous options, we selected a few top models that are both fast enough for our capacity in Google Colab and have code available and are easy to implement.
The selected models for use with the Sino-nom Character Localization dataset include:

- YOLOv5: A very fast and easy-to-use PyTorch model that achieves state-of-the-art (or near state-of-the-art) results.
- YOLOv8: YOLOv8 is a state-of-the-art object detection and image segmentation model created by Ultralytics, the developers of YOLOv5.
- Detectron2: Detectron2 has its own model zoo for computer vision models written in PyTorch.
- DETR: Detection Transformer (DETR) is an end-to-end object detection model implemented using the Transformer architecture.
-

2.2 Assigning Tasks to Team Members

To ensure efficient progress, we have broken down the project into manageable tasks and milestones. These tasks have been carefully assigned to team members based on their strengths and expertise. Regular communication and collaboration are maintained to ensure smooth coordination and progress.

- Van Anh:
  - Sino-nom data analysis.
  - Implementing a pipeline in DETR.
  - Evaluate output result in validate test in mAP.
- Duc Anh:

  - Implementing a pipeline in YOLOv8.
  - Evaluate output result in validate test in mAP.
  - Error Analysis

- Huong:
  - Conducting survey of existing models in Object Detection and datasets in Sino-nom Handwriting.
  - Implementing a pipeline in Detectron2.
  - Data Augmentation using this dataset.

- Uyen:
  - Implementing a pipeline in YOLOv5.
  - Evaluate output result in validate test in mAP.
  - Data Augmentation using other dataset.

- All team members:
  - All team members will collaborate to optimize the best model achieved during phase 2. Additionally, we will collectively fine-tune the model with new data to achieve the best possible results.
  - All off us will contribute to documentation and report writing.