**Image Processing: Final Report on Sino-nom Character Localization**

**Author**: Le Minh Duc

**Course**: INT3404E - Image Processing

**Date**: May 29, 2024

**Objective**: The project aims to develop a novel system for localizing Sino-nom characters in images. Sino-nom characters, a combination of Chinese characters and Vietnamese script, pose unique challenges for image processing systems. The study employs YOLOv8 and SSD approaches, concluding that YOLOv8 offers superior performance.

1. **Introduction**:
   * The study focuses on developing a deep learning-based system for localizing Sino-nom characters within images.
   * Previous methods struggled with accurately identifying and localizing these complex characters due to their intricate nature and mixed script.
2. **Methods**:

2.1 Dataset: - The dataset provided by the course’s teaching assistant contains images featuring Sino-nom characters on various backgrounds.

2.2 Localization Models:

* YOLOv8: A state-of-the-art object detection model known for its speed and accuracy. YOLO (You Only Look Once) processes images in real-time and provides bounding boxes around detected objects.
* SSD (Single Shot MultiBox Detector): Another efficient object detection model that divides the image into a grid and predicts bounding boxes and confidence scores for potential objects.

2.3 Training and Implementation: - Both models were trained on the Sino-nom character dataset using annotated images. - The training process involved optimizing the models to accurately predict the location of Sino-nom characters within diverse and complex images.

1. **Experiments**:

3.1 Evaluation Metrics: - Mean Average Precision (mAP): Measures the accuracy of the models in terms of precision and recall at various intersection-over-union (IoU) thresholds. - Inference Time: Assesses the speed at which the models can process and predict character locations in real-time applications.

3.2 Results: - YOLOv8: - Achieved higher mAP scores compared to SSD, indicating better precision and recall in character localization. - Faster inference time, making it suitable for real-time applications. - SSD: - Lower mAP scores and slightly slower inference times compared to YOLOv8. - Still effective but less optimal for the specific task of Sino-nom character localization.

1. **Discussion**:
   * The experiments demonstrated that YOLOv8 significantly outperforms SSD in both accuracy and speed for Sino-nom character localization.
   * YOLOv8’s ability to quickly and accurately detect characters makes it ideal for applications requiring real-time processing.
2. **Future Works**:
   * Data Augmentation: Implement more advanced data augmentation techniques to improve model robustness.
   * Hybrid Models: Explore combining YOLOv8 with other models to enhance localization accuracy further.
   * Transfer Learning: Leverage pre-trained models on larger datasets to improve initial training efficiency and accuracy.
3. **Conclusion**:
   * This report presented the development and evaluation of a localization system for Sino-nom characters using YOLOv8 and SSD models.
   * The findings highlight YOLOv8’s superiority in terms of accuracy and speed, making it the preferred choice for real-time Sino-nom character localization tasks.
   * Future research will focus on further improving the model’s performance through advanced techniques and hybrid approaches.

**References**

1. Redmon, J., Farhadi, A. (2018). YOLOv3: An Incremental Improvement. arXiv:1804.02767 [cs.CV].
2. Liu, W., Anguelov, D., Erhan, D., Szegedy, C., Reed, S., Fu, C.-Y., Berg, A. C. (2016). SSD: Single Shot MultiBox Detector. arXiv:1512.02325 [cs.CV].
3. Bochkovskiy, A., Wang, C.-Y., Liao, H.-Y. M. (2020). YOLOv4: Optimal Speed and Accuracy of Object Detection. arXiv:2004.10934 [cs.CV].