Building an Adaptive Vietnamese License Plate Recognition and Retrieval System using Multi-Task Deep Learning

Phuoc Minh Hieu PHAM[1], Sy Sieu CAO[1], Le Phu Trung HUYNH[1*]

[1]University of Management and Technology HCMC

*Corresponding author: trung.huynhlephu@umt.edu.vn

Abstract

Automatic License Plate Recognition (ALPR) is an essential component of intelligent transportation, yet its performance is often significantly degraded by real-world image distortions and regional plate format complexities. This research addresses these challenges by proposing a highly adaptive, multi-task deep learning framework specifically designed for the Vietnamese license plate context. The system targets the unique diversity of Vietnamese plates while robustly handling low-quality image inputs.

The proposed framework operates as a multi-stage, conditional pipeline. First, a real-time object detection model is employed to localize all license plate instances. The core component of the system is a lightweight Quality Assessment Module (QAM), which acts as an intelligent router, analyzing and classifying each detected plate into one of three distinct categories: "clear", "restorable", or "unrestorable". The system's adaptive nature is demonstrated in the subsequent multi-branch routing: "restorable" images are selectively forwarded to a specialized restoration neural network. Conversely, "clear" images bypass this resource-intensive step. Finally, "unrestorable" images are rejected entirely, preventing erroneous processing and optimizing overall system throughput. A robust Optical Character Recognition (OCR) model is then used to transcribe the character string from both "clear" and successfully "restored" plates. Finally, the recognized string is used as a query key for retrieving vehicle information from an associated database. This multi-task approach—integrating detection, quality assessment, conditional restoration, and recognition—demonstrates significant accuracy improvements under challenging real-world conditions compared to traditional, non-adaptive pipelines. The system provides a robust and efficient solution for practical ALPR and information retrieval applications within the specific context of Vietnam.

Keywords— Vietnamese License plate recognition, multi-task learning, computer vision.

I. INTRODUCTION

Automatic License Plate Recognition (ALPR) serves as a cornerstone technology within modern Intelligent Transportation Systems (ITS). Its applications are extensive and critical, underpinning systems for automated toll collection, traffic law enforcement, smart parking management, and vehicle access control. The efficacy of these applications hinges on the system's ability to provide accurate and real-time vehicle identification.

However, the performance of conventional ALPR systems degrades significantly when deployed in unconstrained real-world environments. These systems often fail when faced with a wide spectrum of image quality issues, including motion blur from fast-moving vehicles, severe glare from headlights or sunlight, low-light noise, and geometric distortions from oblique camera angles.

This research specifically addresses the challenges within the Vietnamese license plate context. This environment presents unique regional complexities, including a high diversity of plate formats, various background colors (e.g., white, blue, yellow), differing character layouts. A "one-size-fits-all" ALPR solution is often insufficient to handle this variability, leading to high error rates.

II. RELATED WORK

The proposed system comprises four sequential yet adaptive stages: (1) license plate detection, (2) image quality assessment and routing, (3) conditional image restoration, and (4) character recognition and retrieval.

A. License plate detection

Despite these advancements, existing detectors assume relatively clean input images and do not adapt to severe degradations prevalent in Vietnamese traffic surveillance systems, including extreme viewpoint angles, specular glare, motion blur, and low-resolution sensors [3]. For instance, while segmentation-based methods on edge devices [1] and OpenALPR adaptations [2] have been proposed for local deployment, they remain vulnerable to real-world distortions. Moreover, none incorporate preprocessing quality assessment to filter or route degraded inputs prior to recognition. This critical limitation motivates our use of YOLOv8-nano—building on the robust real-time detection foundation of YOLO [4]—augmented with a novel Quality Assessment Module (QAM) to enable adaptive downstream processing.

B. Image quality assessment and routing

Although multi-angle models have been explored [3], they do not dynamically assess input quality or selectively restore degraded plates—leading to cascading errors in recognition.

C. Conditional image restoration

D. Character recognition and retrieval

III.   METHODOLOGY

IV.   CONCLUSION

V.   REFERENCES45

References

[1] Duy Dieu Nguyen, Tan Sang Vo, Manh Hung Le, and Minh Son Nguyen. *An integration of segmentation technique on edge devices for license plate recognition.* VJSTE, vol. 67, no. 1, pp. 3-13, Mar. 2025.

[2] H. Tran, G. Ma, T. Nguyen, and T. Cao. *Building Vietnam's License Plate Recognition System Based on OpenALPR.* International Journal of Multidisciplinary Research and Publications, ISSN (Online): 2581-6187, 2023.

[3] D. Tran-Anh, K. L. Tran, and H.-N. Vu. *License Plate Recognition Based on Multi-Angle View Model.* Preprint, 2023.

[4] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi. *You Only Look Once: Unified, Real-Time Object Detection.* Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2016.

[5] Y. Topuz, M. T. Gökcan, S. Yıldız, and S. Varlı. *A Single Detect Focused YOLO Framework for Robust Mitotic Figure Detection.* MIDOG Workshop (workshop presentation), 2025.

[6] L. Agnolucci, L. Galteri, M. Bertini, and A. Del Bimbo. *Restoration of Analog Videos Using Swin-UNet.* Proceedings of the ACM International Conference on Multimedia (ACM MM), 2022.

[7] M. V. Conde, U.-J. Choi, M. Burchi, and R. Timofte. *Swin2SR: SwinV2 Transformer for Compressed Image Super-Resolution and Restoration.* ECCV Workshops / LNCS, 2022.

[8] B. Shi, X. Bai, and C. Yao. *An End-to-End Trainable Neural Network for Image-based Sequence Recognition and Its Application to Scene Text Recognition.* IEEE Transactions on Pattern Analysis and Machine Intelligence (TPAMI), 2017.

[9] R. Baena, S. Kalleli, and M. Aubry. *General Detection-based Text Line Recognition.* Preprint, 2024.

[10] D. Bautista and R. Atienza. *Scene Text Recognition with Permuted Autoregressive Sequence Models.* In Proceedings of the European Conference on Computer Vision (ECCV), 2022.

[11] X. Wang, Y. Li, H. Zhang, and Y. Shan. *GFP-GAN: Towards Real-World Blind Face Restoration with Generative Facial Prior.* Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2021.

[12] R. Suvorov, E. Logacheva, A. Mashikhin, A. Remizova, A. Ashukha, A. Silvestrov, N. Kong, H. Goka, K. Park, and V. Lempitsky. *Resolution-robust Large Mask Inpainting with Fourier Convolutions.* Proceedings of the IEEE Winter Conference on Applications of Computer Vision (WACV), 2022.

[13] H. Zhou et al. *MaskNet: Towards General Instance Segmentation with Mask Conditioning.* Proceedings of the European Conference on Computer Vision (ECCV), 2020.