

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

Phuoc Minh Hieu PHAM¹, Sy Sieu CAO¹, Le Phu Trung HUYNH^{1*}

¹University 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.

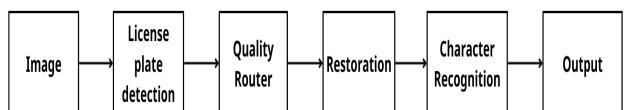


Figure 1: Overview of the proposed adaptive ALPR pipeline.

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 [1]. For instance, while segmentation-based methods on edge devices [2] and OpenALPR adaptations [3] have been proposed for local deployment, they remain vulnerable to real-world distortions. Moreover, none incorporate pre-processing 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

Assessing the quality of detected license plate images is crucial for robust ALPR under real-world degradations. No-reference image quality assessment (NR-IQA) methods have gained prominence due to their applicability without pristine references. Early deep learning approaches leveraged CNNs to predict perceptual quality scores [5, 6], while recent works incorporate geometric priors [7] or lightweight architectures for efficiency [8]. These methods, however, are primarily designed for *general image aesthetics or distortion classification*, not for *task-specific routing* in ALPR pipelines.

To enable adaptive processing, we require a lightweight classifier capable of categorizing plate images into “clear”, “restorable”, or “unrestorable”. Among candidate backbones—ResNet50, MobileNetV2, and EfficientNet-B0 [9]—*EfficientNet-B0* was selected due to its superior parameter-to-accuracy ratio and proven efficacy in no-reference contrast and distortion assessment. Fine-tuned on a synthetic dataset of Vietnamese license plates with controlled blur, noise, and compression artifacts, our *Quality Assessment Module (QAM)* dynamically routes inputs: clear plates proceed directly to OCR,

recoverable plates will be moved to the recovery module, and unrestorable cases are rejected to prevent error propagation.

Although multi-angle detection models have been proposed [1], *none integrate real-time quality evaluation with conditional restoration*, resulting in cascading recognition failures under adverse imaging conditions. Our multi-branch design addresses this gap by optimizing both accuracy and computational efficiency.

C. Conditional image restoration

Restoring degraded license plate images is essential for improving OCR accuracy in low-quality inputs. Classical techniques, such as noise modeling and blur estimation [10], rely on hand-crafted priors and perform poorly under complex, compound distortions. Traditional machine learning approaches [11] offer marginal improvements but lack generalization. The advent of deep learning has introduced more powerful paradigms: convolutional neural networks (CNNs) [12, 11] capture local patterns effectively but suffer from limited receptive fields; *Transformer*-based models [13, 14, 15, 11] excel at modeling long-range dependencies through self-attention, though vanilla Vision Transformers (ViT) incur prohibitive computational cost on high-resolution inputs. Advanced generative methods, including GANs and diffusion models [11], achieve state-of-the-art perceptual quality but risk introducing artifacts (e.g., hallucinated characters) and demand significant resources—unsuitable for real-time ALPR.

For conditional restoration in resource-constrained environments, lightweight efficiency and fidelity are paramount. Among *Transformer*-based restoration backbones, *Swin2SR* [15] stands out by combining the hierarchical *Swin Transformer V2* with compressed image super-resolution objectives, enabling efficient processing of low-resolution, blurry, or noisy inputs while preserving textual integrity. Fine-tuned on synthetically degraded Vietnamese license plates, *Swin2SR* is selectively applied only to “restorable” images identified by the QAM, bypassing clear plates to minimize latency and preventing artifact-induced OCR failures in unrestorable cases.

Despite extensive progress in general image restoration [11], no prior ALPR system integrates task-aware quality routing with conditional, text-preserving restoration. This gap underscores the novelty of our adaptive pipeline.

D. Character recognition and retrieval

III. METHODOLOGY

IV. CONCLUSION

V. REFERENCES

References

- [1] D. Tran-Anh, K. L. Tran, and H.-N. Vu, “License plate recognition based on multi-angle view model,” Preprint, 2023.
- [2] D. D. Nguyen, T. S. Vo, M. H. Le, and M. S. Nguyen, “An integration of segmentation technique on edge devices for license plate recognition,” *VJSTE*, vol. 67, no. 1, pp. 3–13, March 2025.
- [3] 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*, 2023.
- [4] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, “You only look once: Unified, real-time object detection,” in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2016.
- [5] S. Bianco, L. Celona, P. Napoletano, and R. Schettini, “On the use of deep learning for blind image quality assessment,” *arXiv preprint arXiv:1602.05531*, April 2017.
- [6] H. Talebi and P. Milanfar, “Nima: Neural image assessment,” *arXiv preprint arXiv:1709.05424*, April 2018.
- [7] N.-H. Shin, S.-H. Lee, and C.-S. Kim, “Blind image quality assessment based on geometric order learning,” pp. 12 799–12 808, 2024.
- [8] N. B. Nasim Jamshidi Avanaki, Abhijay Ghildyal and S. Zadtootaghaj, “Lar-iqa: A lightweight, accurate, and robust no-reference image quality assessment model,” *arXiv preprint arXiv:2408.17057v2*, September 2024.
- [9] J. H. Joloudari, B. Mesbahzadeh, O. Zare, E. Arslan, R. Alizadehsani, and H. Moosaei, “No-reference image contrast assessment with customized efficientnet-b0,” *arXiv preprint arXiv:2509.21967*, 2025.
- [10] M. Maru and M. C. Parikh, “Image restoration techniques: A survey,” February 2017.
- [11] P. K. H. S. A. T. Nikita Singhal, Anup Kadam and Pranay, “Study of recent image restoration techniques: A comprehensive survey,” June 2025.
- [12] Y. L. E. B. M. D. Peng Liu, Xiaoxiao Zhou and R. Fang, “Image restoration using deep regulated convolutional networks,” *arXiv preprint arXiv:1910.08853v2*, June 2024.
- [13] J. Liang, J. Cao, G. Sun, K. Zhang, L. Van Gool, and R. Timofte, “Swinir: Image restoration using swin transformer,” *arXiv preprint arXiv:2108.10257*, August 2021.
- [14] L. Agnolucci, L. Galteri, M. Bertini, and A. D. Bimbo, “Restoration of analog videos using swin-unet,” in *Proceedings of the ACM International Conference on Multimedia (ACM MM)*, 2022.
- [15] M. V. Conde, U.-J. Choi, M. Burchi, and R. Timofte, “Swin2SR: Swinv2 transformer for compressed image super-resolution and restoration,” in *Proceedings of the European Conference on Computer Vision (ECCV) Workshops*, 2022.